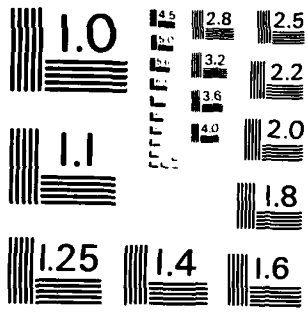
AD-A132 054 THE DEVELOPMENT OF SELECTION STANDARDS FOR THREE NAVY RATINGS WHICH VARY IN LEVEL OF COMPLEXITY(U) NAVAL POSTGRADUATE SCHOOL MONTEREY CA K W NESBITT JUN 83 1/2

UNCLASSIFIED F/G 5/9 NL



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS - 1963 - A

NAVAL POSTGRADUATE SCHOOL

Monterey, California

THESIS

THE DEVELOPMENT OF SELECTION STANDARDS
FOR THREE NAVY RATINGS WHICH VARY
IN LEVEL OF COMPLEXITY

by

Kelvin W. Nesbitt

June 1983

Thesis Advisor: Richard S. Elster

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER	2. GOVT ACCESSION NO. A132057	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) The Development of Selection Standards for Three Navy Ratings which Vary in Level of Complexity		5. TYPE OF REPORT & PERIOD COVERED Master's Thesis June 1983
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Kelvin W. Nesbitt		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Naval Postgraduate School Monterey, California 93940		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Naval Postgraduate School Monterey, California 93940		12. REPORT DATE June 1983
		13. NUMBER OF PAGES 141
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Selection Standards, Enlistment Standards, Utility Analysis, Cross Validation, Navy Ratings, Personnel Selection, Enlisted Personnel, Navy Cohort.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This thesis is concerned with selection standards for three U.S. Navy ratings which vary in terms of their complexity. The relevant literature is reviewed and a general selection standards approach is developed. This approach is then applied to sub-samples of a large U.S. Navy cohort of enlistees who all had the opportunity of serving for at least four years. Within each rating, prediction equations are developed which link various

DD FORM 1473, 1 JAN 73 EDITION OF 1 NOV 65 IS OBSOLETE

S/N 0102-LF-014-6601

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

Block 20 (continued)

data available prior to the beginning of the enlistee's service with three criterion measures of performance. Analyses are performed separately for groupings within ratings by, race and sex. Utility analysis is employed to help determine optimum cutoffs on predictors. Many potentially useful predictive relationships are found and amongst the results is the finding that for some ratings, ability subtests are negatively related to criteria of performance. Other results are discussed and recommendations regarding implementation and future research are made.

Approved for public release; distribution unlimited.

The Development of Selection Standards for Three Navy Ratings which Vary in Level of Complexity

by

Kelvin W. Nesbitt
Major, Psychology Corps, Australian Army
B.S., University of Newcastle, 1969
Diploma in Applied Psychology, University of Adelaide, 1978

Submitted in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE IN MANAGEMENT

from the

NAVAL POSTGRADUATE SCHOOL
June 1983

Author: ______________________________

Approved by: ______________________________
Thesis Advisor

____
Second Reader

____
Chairman, Department of Administrative Sciences

____
Dean of Information and Policy Sciences

ABSTRACT

This thesis is concerned with selection standards for three US Navy ratings which vary in terms of their complexity. The relevant literature is reviewed and a general selection standards approach is developed. This approach is then applied to subsamples of a large US Navy cohort of enlistees who all had the oppurtunity of serving for at least four years. Within each rating, prediction equations are developed which link various data available prior to the beginning of the enlistee's service with three criterion measures of performance. Analyses are performed separately for groupings within ratings by race and sex. Utility analysis is employed to help determine optimum cut-offs on predictors. Many potentially useful predictive relationships are found and amongst the results is the finding that for some ratings, ability subtests are negatively related to criteria of performance. Other results are discussed and recommendations regarding implementation and future research are made.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

I. INTRODUCTION

The Armed Forces in many countries have been ardent users of the psychological techniques of selection and classification. The Forces' involvement with these procedures has been for such a considerable period of time that, in fact, it could be argued that the development and refinement of these techniques has been greatly enhanced by the manpower demands of the military. In war, and in preparation for war (i.e. mobilization), the military required techniques which allowed the assesssment of the training capabilities of prospective civilians for their rapid transition from civilian to uniformed status. It was inefficient to have every potential enlistee embark on a course of training when psychometric tests could be used to exclude those individuals with the lowest chances of success. Since the Second World War, the peace time forces, largely through the development of more complex weapons systems, have come to demand much more specialized manpower skills. The Forces' requirement for selection and classification, therefore, has not only involved making military personnel out of civilians (i.e. selection), but also dealt with the correct placement of enlistees by military occupation (i.e. classification). Both the size and the diversity of the Armed Forces' manpower requirements have make it a very fertile testing ground for the application of these psychological procedures.

The present thesis is concerned with the refinement of selection standards for three employment categories (ratings) in the United States Navy. The three ratings have been chosen on the basis of the complexity of the tasks in the ratings. In increasing order of complexity they are:

Ship's Serviceman, Personnelman and Aviation Technician. All three ratings have currently applied test score entry standards based on the Armed Services Vocational Aptitude Battery (ASVAB). The aim is to examine the possibility of improving the selection criteria for each rating, taking into account rating specific utilities of correct and incorrect selection decisions. At the same time, the selection standards across jobs will be compared. Such an inter-job analysis will illustrate the relationship, if any, that job complexity has with the selection standards. This latter analysis will be qualitative (what kinds of predictors are more suited to what type of job) and quantitative (how does the level of accuracy of selection vary with job complexity).

Selection standards research for the US Navy is important for a large number of reasons. From the point of view of manpower planning, better selection standards may mean smaller training attrition, better job performance, longer careers and more rapid promotion. These outcomes have the effect of lifting the burden of maintaining the Navy's end strength from new enlistments by having greater numbers in the career force. This has even more significance with the projected decline in the services' traditional manpower pool over the next decade, particularly in that when manpower resources are scarcer, the disutility of rejecting potentially successful applicants is much greater than it is in the current recruiting climate. In a broader context, current government regulation explicitly forbids selection practices which may have adverse impacts on minority or other groups in the community (Uniform Guidelines on Employee Selection Procedures, 1978) without first investigating the use of "alternative selection procedures that are equally valid, but that produce less adverse impact" (Reilly and Chao, 1982). It is encumbent upon all employers, including the US Navy, first to determine the validity of their current procedures and then to examine the possibility of using alternative methods which may either improve on these standards or are equally valid standards which have less adverse impact. Thus there is a strong extrinsically initiated incentive for an organization to be interested in, and to be able to measure the validity of, its own selection procedures.

There is a wealth of research on selection standards both in the US Navy and in many other employment settings. Given the aim of looking at selection, in relation to the complexity of the job and the validity of the procedures, the next chapter is devoted to literature surveys. The first section deals with the theoretical aspects of selection research, while the second is a review of the findings of previous research regarding those factors which seem most predictive of job success, in general, and for the services in particular.

The third chapter brings the information of the previous one together so as to make a clear statement of the purpose, aim, design and expected outcome of the research.

Chapter four deals with the method. First, the subject ratings are described. This includes a description of each rating, the standards currently applied and an examination of the paths through which individuals enter the ratings. It will include the complexity grading of each rating. Next comes a description of the data base for this research: where it comes from, what is included and an assessment of the quality of the data. In the next section, the personnel are described. This is followed by a section which concerns the criterion measures and predictors: What are they?, and, How they are measured? Another section describes the estimates of the utility of correct predictions and the disutilities associated with misclassifications. Next, the cross-validation procedures used for the estimation of the validity coefficient are described. Finally, the statistical techniques are discussed.

Chapter five concerns results. The first section deals with descriptive statistics for each of the three ratings, which includes age distributions, racial and sex mixes as well as prior education levels. The next section concerns the results from using regression techniques. The final section presents results in which the direct effects of selection cut-offs are examined with respect to maximising total expected return for the selected procedure.

Chapter six is for the discussion of the results. It is divided into two sections: expected outcomes and other findings. Particular attention is paid to across rating comparisons.

Chapter seven provides a summary and a statement of the conclusions and recommendations.

II. LITERATURE REVIEW

The purpose of this chapter is to review some of the literature relevant to selection standards research. The first section deals with a general overview of the technical and psychological aspects: How should selection standards be established? What is the concept of validity and how does it apply? What types of experimental design are possible and appropriate in this type of research? Then, in more detail specific theoretical issues are addressed. These are job analysis, criterion selection, predictor selection, validation and cross validation. The second section reviews research conducted specifically in selection and/or attrition. These questions are addressed: What variables have been found to be predictive of success in the work environment? Specifically, what variables are associated with success in a service environment? And finally, what factors predict success in a Navy environment? The final section provides a summary of the theory of, and the previous research in, selection.

A. THEORETICAL LITERATURE

1. Classical Model

A classical model for establishing selection standards is suggested by Drenth (1971). Like most classical approaches it is probably never rigidly applied, however Drenth's suggestion gives a good description of the steps involved and is a useful starting point for this section. These steps are summarised in Table I.

TABLE I

Classical Model of Selection

Step 1.	Job Analysis
Step 2.	Choice of a Criterion
Step 3.	Try-out Series of Tests
Step 4.	Testing of Experimental Group
Step 5.	Validation of Tests
Step 6.	Cross Validation
Step 7.	Tests in Final Form

Source: Drenth, P "Theory and Methods of Selection", 1971.

The first task is a job analysis in order that the psychological requirements of the job can be understood. In this step it is necessary to convert the description of the job into specific factors or variables which are defined and, hopefully, are measurable. The second task is to determine the criterion which is the operational means of measuring an employee's job performance. A knowledge of the criterion value should permit a ready differentiation of the good performers from the poorer ones. Ideally, it will be a metric or ordinal scale which will indicate ranges or grades of performance. In the third step, the job analysis is used to make an 'educated' guess as to what kinds of tests or other factors are important for predicting success in the job for which the standards are designed. The preceding steps are necessary precursors to the next group which deal with the empirical portion of the process. The fourth step consists of the try-out (or pilot) tests which are administered to a representative sample of the population for which the standards are to be developed. Scores or results are recorded and retained while the employees are given a period of time to perform on-the-job. In step five, these subjects are measured on the performance criteria developed in step two, and the results are correlated with the test scores gathered in the previous step. From this analysis the merit of each test is determined leading to the a choice of tests most useful for selection and their appropriate cut off scores. Step six is critical: the cross validation of the selection standards for new group of subjects. In this step the same procedure as applied in steps three to five is repeated using the chosen selection tests. Step seven is largely administrative. The selection tests are set up in a standard form. For small organisations this step may not be essential, but for larger ones the standardization of the tests is highly desirable, especially where there may be several different testing stations.

The Drenth classical model is a very simplistic scheme of the procedure for establishing selection standards. For one thing, it appears to assume that the organization has no selection procedures in existence, simply allowing individuals to drift into occupations. While it may not be strictly formalised, the organization probably has a good notion of what is required in the job and applies this information as a screening device. Despite these limitations, the model alludes to some very important issues which are discussed in the subsequent paragraphs.

2. Validity and Research Design

Validity is a very important issue. In the simple case above, it is described as some means of comparing the selection test results with subsequent performance in order to validate selection standards. Validity answers the question: do these test scores tell us a sufficient amount about the individuals' subsequent performance so that, based only on a knowledge of their selection test results, one is able to choose the better performers. Statistically, it is usually repesented by a correlation coefficient which expresses the degree of association between the selection test scores and the measure of subsequent performance.

Two main types of research designs have been defined with respect to selection validation. These are predictive and concurrent validity (see Tiffin and McCormack, 1970). Predictive validity is that implied in the model above: selection tests are given to a group of applicants and then at some later date, criterion measures of performance are made. The test scores and criterion measures are then compared. For concurrent validity design, the tests scores and criterion measures are gathered virtually at the same time. For both types, the tests, or a subset of the tests, are considered valid, if the correlations are large enough. Until recently, the predictive validity design was considered far superior than concurrent validity approach (Drenth, 1971). However Barrett et al (1981) persuasively argue that the evidence does not support this position. They suggest that a large number of independent studies, including a recent meta analysis of 537 separate studies, have indicated virtually no difference between the magnitude of the validities found for concurrent and predictive designs. They also cite one researcher who specifically compared the two designs in a national validation of selection procedures for male transit bus drivers and found that the validity coefficients were comparable. These arguments are convincing but what are the significant differences between the two validation designs?

Guion and Cranny (1982) suggest that the key difference is that for predictive designs there is a lapse of time between test and criterion measurement whereas in concurrent designs the measurements are made simultaneously. However, there is a second fundamental difference between the two:

predictive designs invariably test groups of applicants, rather than employees, as tested in concurrent designs. There are some problems for concurrent designs in relation to the restriction of range of tests scores. The employees who are the subject of concurrent designs have been preselected contributing to a narrowing of the range of data. As is noted by Hammer and Landau (1981, p. 576), range restriction shrinks the feasible range of the correlation coefficient; that is, its feasible maximum value becomes less than one. Predictive designs are not completely free from correlation problems, since there is a time difference between tests and criteria. Scores on the latter may well be due to factors other than those implied by the tests scores. Intervening variables (e.g. training factors, motivational changes) could also inhibit criterion performances. Additionally, attrition may reduce the size of the group so that there are fewer individuals available for criterion measurement than were orginially tested. It can be seen that both concurrent and predictive designs are susceptible to shrinkage of validity coefficients.

While it seems only logical for an organization, particularly a large one, to have valid selection procedures, government regulation has a significant impact on these procedures requiring employers to have selection systems which have been tested through a general validity model. Hsu (1982) points out that the Uniform Guidlines on Employee Selection Procedures (1978) specify "that a necessary condition for use of a selection procedure, when that procedure produces adverse impacts on the hiring rates of certain minority groups, is that this procedure be demonstrated valid" (Hsu 1982, p. 509), but that this is not a sufficient condition for its use. Hsu explains that for an organisation to use any existing selection procedure that produces adverse minority impacts, it must have higher validity than alternatives which have no adverse impacts. Many civil actions have been won during the last decade by employees (see Rowland *et al*, 1980, pp. 64-85) against employers who have had selection systems for which the validity has been unknown or inadequately established. So there is a considerable amount of external pressure for the organization to, first, know what the validity of its procedure is, and to, second, establish, in the case that there are adverse impacts, that there is no alternative system which is equally valid that will produce a less adverse impact.

Given this general overview of the technical literature, it is now possible to deal in more detail with other issues.

3. Job Analysis

It is probably a fairly safe assumption that very few attempts at introducing selection standards begin with a thorough job analysis (such as is described in Dunnette, 1976, Chapter 15). This is so for a variety of reasons, not the least of which is that it is very time consuming and requires a considerable amount of effort with, at best, very little by way of quickly identifiable return for these labours. A job analysis is not part of this research, but Chapter IV includes a brief job description of the three Navy ratings involved, in order to put them into an appropriate context. Presumably thorough job analyses have been done for these ratings and used for the development of the selection procedures.

4. Criterion Selection

Dunnette's book ("The Handbook of Industrial and Organizational Psychology", 1976) has two excellent chapters on the selection of an appropriate criterion, which is the second step in Table I. This is perhaps the hardest step of all and has been one of the "key problems in industrial and organizational psychology, as evidenced by the massive efforts designed to clarify its theory and improve its measurement" (Smith, 1976, p. 745). Smith specifies three requirements of a criterion of performance:

a. that it be relevant of some organizational or individual goal, but neither biased nor trivial;

b. that it be reliable and "involves agreement between different evaluations, at different periods of time and with different although apparently similar measures" (Smith, p. 746);

c. that it be practical: that is, available, plaus--ible and acceptable to those who will use it for decisions.

Another issue that has received much research is the choice between a single performance criterion and multiple criteria. In the classical model, the emphasis was on a single criterion which met all the requirements mentioned above (Guion, 1976, p. 783), but the weight of recent opinion is towards multiple criteria (Smith, 1976, p. 747). The first argument in favour of the multiple criteria solution is on logical grounds: the various possible criteria are all different ways of measuring the same unitary concept and using more than one measure gives a better estimate of this concept's true value. It could be that two individuals with equivalent total performances could achieve these results through quite different behav-iour patterns. The other arguement for multiple criteria is empirical: "an overwhelming majority of studies involving statistical analyses of sets of criterion measures finds that these analyses rarely yield a single general factor...several criterion measures are necessary to account for the variance in a criterion correlation matrix" (Smith, 1976, p. 748). The choice of criteria and how to combine them are difficult processes and must all too often be based on convenience rather than their relevance to long term effectiveness.

Several recent studies have proposed the use of utility concepts in performance measurement which could be used as a criterion (Landy, Farr and Jacobs, 1982; Schmidt and Hunter, 1981). Utility concepts include such things as the number of successful employees hired, increases in the average level of performance, and the amount of savings in dollars and cents, as a means of evaluating the effects of various selection strategies. Most recent applications of utility concepts deal with the value to the organisation of the hired individual which was derived by estimating the average value and the standard deviation of performance expressed in dollars. These criteria as they stand are inappropriate for the present study, but the application of utility analysis is discussed below in relation to the determination of optimum cutting scores on predictors.

In addition, Smith (1976) suggests several other variables which have been found useful as providing "hard" and "soft" criterion measures. Under the hard category are tardiness, absences, accidents, tenure or turnover, sales, production, job level and promotions. Even these may contain some soft elements largely due to the effects of human judgements. The soft criteria are those which involve the use of rating scales, usually completed by supervisors. There are a large number of sources of rating scale error, and an even larger number of methods to minimise their effects (see Landy and Farr, 1979; Smith, 1976, pp. 757-764), but these will not be considered in detail here.

5. Predictor Selection

In trying to select tests which may serve as predictors, Guion (1976) suggests three criteria:

a. complex behaviour (i.e. the criterion) cannot be fully predicted by simple means, thus a single predictor is usually inappropriate;

b. complex performance is a function of the individual, but only to a certain extent, therefore in thinking of predictors, it is pertinent to go beyond individual traits and include situational and demographic variables as sources of potential predictors; and

c. complex behavior is not likely to be optimally predicted in precisely the same way for all people, while it is not feasible to have prediction equations for each individual, it may be possible to develop different equations for logically different subgroups.

He also suggests many publications to consult (e.g. Ghiselli, 1966) which provide a summary of predictive validities for various job categories, by type of predictor and by type of criterion. They are, he suggests, a good starting point.

6. Validation

To this point the concept of validity has been described as a correlation coefficient which expresses the relationship between predictor and criterion measures. However, this concept can be extended. As noted by Campbell (1976), the correlation coefficient approach to validity implies:

a. " there is one normally distributed criterion or criterion composite;

b. there is one normally distributed predictor or predictor composite; (and)

c. the relationship between them is linear and homoscedastic." (p. 205)

These are stringent conditions and are often violated. He further challenges the practical use of a validity coefficient: how are decision makers to use say, the value of .55, in determining which of a group of applicants should be selected? The coefficient says little about how to produce the desired outcome. Provided the above three conditions are met, the validity coefficient is useful for determining which of a group of predictors is more appropriate for selection but Campbell describes what he calls "decision centered" validity as a more useful approach.

With a decision orientation, selection research is viewed an attempt to predict discrete outcomes on a criterion which is categorical in nature. The categories to be predicted are made up of those who succeed against those who fail. The task for the predictor then is to establish a "cutting" score which maximises the proportion of correct predictions. But how is the proportion of correct predictions to be defined? In the equations below, the figures A, B, C, and D all refer to the numbers of personnel by predictor/ criterion outcome. A is the number who are predicted to pass the criterion and who do, in fact, pass. B is the number who are correctly predicted to fail the criterion. C is the number who are predicted to pass but who subsequently fail the criterion. Finally, D is the number who are predicted to fail but who would have passed. Using these figures, two examples of how the proportion of correct classifications may be defined are as follows. One method, Equation 2.1, is to take the proportion of correct classifications over the total number to classify. Another approach, Equation 2.2, takes as the statistic the number of successes over the total number who pass. There could be other possible definitions. However for both equations, the idea is to select a "cutting" score on the predictor which maximizes the value of the expression. What is forgotten in both these definitions is that each embodies a value judgement in the way the statistic is defined. Equation 2.1

$$CP = (A + B)/(A + B + C + D) \qquad \text{(eqn 2.1)}$$

$$CP = (A)/(A + C) \qquad \text{(eqn 2.2)}$$

applies an equal weight to both correct prediction outcomes, and does not distinguish between the two types of error (categories C and D in the equations) in terms of their (negative) value to the organisation. Equation 2.2 ignores completely the outcomes for individuals who are not in the predicted success group, in effect assigning a value of zero to each of these outcomes. The question arises then as to what values should be assigned to each outcome for the four possibilities, given in the equations, to derive an appropriate cutting score?

The answer lies in the application of the concept of utility which was referred to briefly above. As described by Campbell (1976), utilities are estimated for each outcome in Equations 2.1 and 2.2. For simplicity in the following discussion, the terms pass and fail refer to criterion results whereas positive and negative refer to predictor score results. Thus, outcome A is a pass/positive outcome: these members are correctly predicted to pass on the criterion. An example of the application of utilities and how they are maximised by varying the cutting score is given

TABLE II

Total Expected Utility for a Predictor Cut-off Value

A. Fixed relative utilities for each outcome

	Predict Failure	Predict Success
Success	-5	10
Failure	10	-30

B. Probability of each outcome for a certain cut-off

	Predict Failure	Predict Success	
Success	.16	.24	.40
Failure	.48	.12	.60

C. Products of utilities and probabilities

	Predict Failure	Predict Success
Success	-.80	2.4
Failure	4.8	-3.6

Total expected utility for this cut-off = 2.80

in Table II (this example is very similar to that given by Campbell, 1976, p. 212). As can be seen from Table II, part A, the pass/positive outcome is given a value of 10 units of utility which is the same value as the fail/negative outcome. This implies that it is as valuable to the organisation to predict a success as it is to correctly predict a failure. The negative signs in the other two cells suggest that each outcome is a disutility to the organisation. Further, it is clear that the relative values in these cells indicate that the disutility of incorrectly identifying a failure, is far greater than incorrectly identifying an individual who would have been successful. In fact the former disutility is six times as great and it is not difficult to see why this would be so. Allowing a future failing individual in to the organisation is much more expensive because of the larger amount of investment which is lost (e.g. training expenses, salary) compared with the relatively small investment lost on the rejected applicant who would have been a success. The question remains however as to how to use these utilities in the deriving selection cut-offs.

To see how the utility values are used for determining the effectiveness of a selection procedure, it is necessary to examine all three parts of Table II. Part B shows the proportion of the sample residing in each cell based upon a cut-off score on the predictor. It forecasts that 38% of the applicants will be successful. Treating these cell entries as a probability of belonging to that cell, it is now possible to calculate the expected utility for each cell by multiplying the probability by the corresponding utility. The result is shown in part C of the table. For a cutting point of 38% predicted successes, the total expected utility is simply the sum of each of these cells: 2.8 is the expected utility for this particular cutting score. Using this same procedure the next possible cut-off score on the predictor can be evaluated. Clearly, this cut-off will most probably produce different proportions of cases in each cell, which will in turn change the value of the total expected utility. Continuing this process for all possible cut-offs produces a total expected utility for each cut-off level. The cut-off which should be chosen is the one that maximises the expected utility. Of course if all cut-off levels result in negative expected total utility, then the selection device is of no use to the organisation and it should not be used.

Within the US Navy setting, it is possible to make estimations of the various utilities associated with each of the four possible predictor/criterion outcomes through cost and productivity analyses by rating. Utilities for three of the four possible outcomes appear to be relatively straightforward, provided some simplifying assumptions are made. These three (associated with A, B and C in the equations) may be estimated from the Billet Cost Model (Assessment Group, 1983). However, the disutility of the other cell, misclassified pass (D in the equations) is more difficult. For the defense force the disutility of this outcome is very much a function of the recruiting market. During lean recruiting times, when quotas are hard to fill, to reject a potentially successful candidate has greater disutiltiy than it would in the current recruiting climate. The utilities associated with the other three cells should be impervious to such fluctuations. The actual derivation of utilities is discussed in the Method chapter.

7. Cross Validation

Depending on the statistical technique employed, the results of selection standards research will be effected by certain characteristics specific to the sample used in the study (Weiss, 1976, p. 332). This means that the results will reflect not only the true relationships between the criterion and predictor variables, but they will also reflect relationships which are unique to that particular sample, and this uniqueness may not be generalised to the population. To avoid these problems selection results are usually cross-validated.

It is not necessary to cross-validate in the same manner which was described for the classical selection model. This would involve the complete replication of the original research on a new group of subjects. More efficient and less costly methods are available. Campbell (1976, p. 214) describes what he calls the 'double cross-validation' design. Here the sample is randomly split into two equally sized subsamples. The statistical techniques are applied independently to each subsample to derive the decision rules for the final version of the selection device. The decision rules for subsample 1 are applied to subsample 2 and the validity estimated. The decision rules for subsample 2 are applied to subsample 1 and another coefficient is calculated. The average of these two coefficients gives the validity for the procedures. Ideally, of course, the decision rules for each subsample will be essentially the same. Marked deviation between subsamples indicates that either one or both results are sample specific and therefore they would be suspect.

The double cross-validation design could apply equally well to the utility method for estimating the optimal cut-off values. Cut-offs could be determined on predictors for one group and then compared with the cutoffs derived from these same predictors on the other subsample.

B. RESULTS FROM PREVIOUS RESEARCH

As was indicated earlier, there is a great deal of research literature on personnel selection within organisations. This includes studies for almost every conceiveable employment type both across the community and, within and by, the services. The purpose of this section is to include sufficient of the previous research so as to get a basis for formulating research hypotheses in the next chapter. It is neither intended nor claimed that what follows is an all inclusive review of personnel selection research.

1. General Results

The history of mental testing is intertwined with that of personnel selection in organisations. During the 1950's and 1960's, many organisations adopted the IQ as a standard measurement for use in selection and placement in virtually any employment setting. But the whole area of human ability testing, including the concept of IQ, came under fire in the late 1960's. The civil rights movement had a lot to do with this, as articles and monographs (e.g. Jensen, 1973) indicated that there were significant minority score differences on these tests. Further, there were claims that minority groups were disadvantaged, through test bias, by organisations which employed psychological tests in their selection procedures. In a recent article Carroll and Horn (1981) sought to make a distinction between the science and the application of human ability measurement. They claimed that the science of human abilities was in a "hase of confusion (as evidenced) when ideas about human abilities are regarded as equivalent to ideas about IQ measurements (and) that valid criticism of particular applications of measurements are not, in general, criticisms of the science of ability measurement" (page 103). They argue that there is no particular fault with the scientific aspects of human ability measurement, claiming that 80% of the variance of ability differences is accountable in test measurement, and suggesting that the application of these tests has caused the difficulty.

Schmidt and Hunter (1981) also discuss this issue, but in a more applied sense. They have derived, over the last five years or so, a procedure for evaluating the significance of individual validity studies, in which they say it is sometimes difficult to detect the presence or effect of sampling bias. Specifically, they directly address many of the criticisms of mental testing put forward in the last decade. These include the notions of low utility (selection procedures have little direct bearing on organisational productivity), test unfairness (selection procedures are biassed against minorities) and test invalidity (selection test validities may be situation or job specific and not truely general). Numerous recent studies are referenced in support of the idea that selection procedures do not have low utility. This leads to the conclusion that the potential effects of employment testing on productivity have been underestimated. The authors deal with test unfairness by suggesting that the factors causing low test scores also produce poorer job performance, for both majority and minority groups. They conclude that employment tests do not cause adverse impacts against minorities and they seem to support Jensen (1980), who demonstrated that when minorities are selected on procedures developed on white samples the minority group is favoured (see Jensen, 1980, p. 515). The third issue of test invalidity is similarly argued with the presentation of many studies which support the notion that tests are valid for employment prediction. One relevant study showed that when jobs were grouped according "to complexity of information-processing requirements, the validities of a composite of verbal and quantitative abilities for predicting on-the-job performance varied from .56 for the highest level job to .23 for the lowest" (page 1133). This is an interesting finding considering the emphasis in this thesis on selection for various complexity levels. It suggests, first, that cognitive ability should be valid for predicting job performance and, second, that the validity should increase with job complexity. Smith and Hunter reach three conclusions:

a. professionally developed ability tests are valid predictors of performance both on-the-job and in training for all jobs;

b. ability tests are fair to minority groups as they do not underestimate the expected job performance of minorities; and

c. ability tests used in selection can effect huge labour and cost savings in an organisation.

In contrast to Schmidt and Hunter's paper, which makes a frontal attack on the critics of ability testing, Reilly and Chao (1982) examine the validity and fairness of some alternative selection procedures as is required by government regulation where there are adverse minority effects. Of the eight alternatives, two are relevant for the present study. These are biograhical data and academic performance measures. Biographical data is that type of information usually given by an application for a job, and includes such things as age, number of years schooling, marital status and number of dependents. These data have been successfully used in predictive cross-validated studies, returning validities not as large as for ability tests, but sufficiently close to suggest they could be a feasible alternative. Academic performance measures, such as grade point averages and course results, have not fared so well. Validities based on performance criteria for these items have been either insignificant or extremely low. They are not reccommended as an alternative.

In sum, therefore, it is clear that mental ability tests are relatively valid predictors of job performance in virtually all settings, and that biographical data are also useful and valid as predictors. Higher validities have been found for more complex jobs when compared with jobs of low complexity.

2. Military Research

As was indicated in the introduction, the military has been a very fertile ground for research in, and the application of, selection standards technology. Almost 150 studies on the US military were recently reviewed by Vineberg and Joyner (1982). These occured between 1952 and 1980 and were selected because they were job performance prediction studies. The authors' paper dealt with criterion and predictor variables and the usual technical measure of validity (i.e. the correlation coefficient) associated with different combinations of these variables. This extensive survey article is discussed in the next few paragraphs.

Vineberg and Joyner categorise the criterion variables used in military studies into three types: proficiency, job performance and suitability. Proficiency variables are objective type measures of either job knowledge, task and/or task element performance. Job performance is expressed as either global ratings, job element ratings (both of which are subject to the rating scale errors referred to above), productivity (applied only to Army recruiters) and the grade or skill level obtained. The suitability criteria are made up of such things as length of service, completion of enlistment term, misconduct measures and recommendations for re-enlistment. Within and across these subdivisions the best median validities obtained were for job knowledge (.40), task performance (.31), suitability (.24) and global rating (.15) from a total of 350 validity coeffients. Based on these values, the authors conclude that the task performance and suitability validities are high enough to suggest they would be useful in selecting military personnel.

Many different types of predictors have been used as the independent variables in selection research. The most popular is aptitude (e.g. mental tests), followed closely by biographic and demographic variables. In most aptitude studies at least one other type of predictor was used. Age and education level (expressed as number of years schooling) were the most frequently used biographical variables. When predicting suitability criteria, the variables age, mental ability and education have consistently been found to be significant. The median validities are: education (.36), mental ability (.24) and age (.21). When all three of these variables are taken into account simultaneously, then validities have ranged from .24 to .39. Some evidence is presented which suggests that when some early military experiences are taken into account (e.g. ratings at recruit training), then the validities for prediction range from the .40s to the .60s.

While the authors conclude that the research suggests the usefulness of these predictors to forecast performances on these criteria, they make two relevant points about the relationship of the variables and how the validities may be improved upon (page 22). The first point is similar to one made earlier in this paper: there are likely to be many important differences within and across jobs, and these need to be taken into account when theorising about what kinds of relationships exist. Jobs may differ in terms of job difficulty, the level of effective ceiling of performance and how many of the encumbents reach this level of performance. These kinds of factors effect what the criterion measure should be and how it should be measured. The present study hopes to address specifically the issue of job complexity.

The second point is that predictor and criterion relationships are not likely to be static over time. With experience on-the-job, those predictors which were useful for predicting "early" criteria of performance tend to "wash out". It may be that as training and experience tend to make the differences in technical proficiency smaller and less significant, then predictors are less able to discriminate between the performances. These points tend to empasise the notion that the expected relationships between predictors and criteria might not be found in a sample, but also that there are probably many alternative explanations for insignificant results.

Studies of predicting military performance have been done in other countries, with similar kinds of results. Some of the more recent ones are summarised by Johnson (1982). An earlier paper, quite relevant to the present study, was done by O'Gorman (1972). He studied regular entrants to the Australian Army and found that a total of ten pre-enlistment variables combined in linear multiple regression were useful in predicting an effectiveness measure of performance. The criterion consisted of assigning the value "1" to those who did not complete their service for reasons other than death or illness, "2" to those who completed three years without promotion, and "3" to those who completed three years and who were promoted to corporal or higher. The significant predictor variables included age, marital status, aptitude scores (general ability and clerical ability), results on a psychiatric inventory, birth order and prior service.

More specifically for the current thesis, there has been some recent pilot research (Lurie 1981) about determining selection standards for the Ships Serviceman rating. In this study Lurie attempted to relate pre-entry variables to an improved criterion of Navy success: rather than the survivability of recruits during their first term he attempted to employ a job performance criterion to determine selection standards. His criterion was a combined advancement and survivalability measure, the object being to predict this variable given AFQT score, age, number of

TABLE III

Percentage of SHs scoring below designated AFQT Score

Grade	AFQT =< 20	AFQT =< 50
E-1	54	83
E-2	50	85
E-3	45	78
E-4	34	70
E-5	30	62

Source: Lurie, P.E. "Relating Enlistment Standards to Job Performance: A Pilot Study for Two Navy Ratings" CNA No. 81-0048 of 15 Jan 1981.

primary dependents and years of education for a cohort group of enlistees. His study used a cohort which joined the Navy in 1973, and records of their service were available up until 1977, allowing for advancement to E5. Regression models were used in calculating the relationships between the variables which allowed for the construction of "service status" probabilities. These probabilities show for various amounts of time in service the likelihood that the sailor will be either advanced/demoted to a particular rank, or discharged from the service. Without controlling for education, Lurie's data (see Table III) show an expected relationship between grade and AFQT. While the numbers are a little confusing, it is clear that as rank increases then the proportion of low AFQT scores declines. However, what is not obvious from Table III is the combined effect of high school graduation and AFQT. Lurie found that high school graduate enlistees with higher AFQT scores, have a better chance of advancement to E3 and E4, while the lower AFQT scorers have a higher rate of attrition. For non-high school graduates, the probabilities indicate that recruits with lower AFQT scores outperform those with the higher scores. This rather striking result cannot be attributed to differential attrition rates as these are virtually identical for both low and high AFQT scorers. Non-high school graduate recruits with lower AFQTs advance more quickly to E4 and E5 than do high scorers. On the other hand, high scorers have a much higher chance of being reduced in rank than the low scorers. Lurie suggests that perhaps non-high school graduates with high AFQTs are dissatisfied with their placement as SHs believing they should be better employed. However no data is presented to support this notion. His overall conclusion is that it appears feasible to set enlistment standards for this rating based upon performance criteria. Unfortunately, his recommendations do not suggest any new standards.

C. SUMMARY

1. Theory

It is now possible to give a short summary of the theoretical literature in this chapter. This is given in point form.

- there are both intrinsic and extrinsic pressures for an organisation to be interested and able to validate its selection procedures.
- validation of selection standards can take either the 'predictive' or 'concurrent' research design as they produce equivalent results.
- all selection research should begin with a thorough job analysis so that relevant information is obtained to assist in the development of both criterion and predictor concepts.
- certain principles are known which assist in the

selection of criterion and predictor measures.

- validity coefficients are susceptible to shrinkage due to a number of factors which may be controlled or adjusted for.

- while correlation coefficients are useful in selecting the most relevant predictors from a large number of potential predictors and as a technical statistic of validity, alone they do not indicate precisely what the selection procedures should be.

- providing utilities can be estimated for the various predictor/criterion outcomes, the decision centered validity design appears the most practical and useful to determine predictor cut-offs which maximise utility and which ultimately determine the selection system's overall usefulness.

- the double cross-validation design is a means of overcoming some sources of sample specific bias in selection standards research.

2. Empirical Results

From the review of previous studies the following points may be summarised.

- intra- and inter-job differences (e.g. complexity) are likely to effect the relationship between predictor and criterion measures.

- validity coefficients for more complex jobs are likely to be higher than less complex ones.

- relationships between criterion and predictor variables are not likely to be stable over time as training and experience appear to "wash out" these relationships by reducing differences in criterion performances.

- significant predictor measures found in military studies include age, marital status, education, ability test scores and other biographical variables.
- criterion measures used previously in the military have included the following either singly or as composites: length of service, advancement, recommendations for re-enlistment and various misconduct measures.
- some recent studies (e.g. Lurie, 1981) have emphasised the unusual moderating effect some variables may have on predictors.

III. AIM AND PURPOSE OF THIS RESEARCH

This chapter is designed to make a clear statement of the aim and purpose of the thesis. It is a chance to synthesise the literature survey of the previous chapter into a form which indicates the expected outcome from the analysis of the ratings chosen for study. The general design of the research and a statement of the research hypotheses are also given.

A. PURPOSE OF RESEARCH

The overall purpose of the present research is to analysise the available data on a cohort of entrants to three US Navy ratings in order to test the possibility that the current selection procedures can be improved. The ratings are Ships' Serviceman, Personnelman and Aviation Technician. The study involves the development of selection standards which it is hoped will enhance the total effective performance of future entrants to these ratings.

There are several secondary purposes. Inter-rating differences will be examined in relation to the level of complexity of each rating. To this end, ratings have been chosen deliberately with the intention of having a wide spread in terms of job complexity. The research is intended to examine rating differences with respect to different criteria of performance both for single and composite measures. Finally, it is intended to apply the concepts of utility analysis (that is, applying estimates of the cost and the productivity of sailors) as a means of determining the selection procedure which will yield the maximum pay-off in terms of total expected return to the Navy.

B. AIM

The aim of this thesis is to investigate the possibility of improving the selection standards for three US Navy ratings.

C. DESIGN

In terms of the theoretical literature, the design of the present study is more 'predictive' than it is 'concurrent'. It is not strictly 'predictive' in the classical sense, as it does not include data on all applicants for the US Navy, but rather has information on applicants who were subsequently enlisted. The research is somewhat 'concurrent' in that sailors within a rating are included for analysis provided they met at least one of the three criteria of inclusion discussed in a later section.

After the selection of appropriate predictor and criterion measures, the next feature of the design is the use of double cross-validation coefficients as an overall gauge of the usefulness of the relationship between predictors and criteria as a selection tool. This involves the random splitting of the total sample into two halves, developing a linear regression relationship for each half, then testing the correlation between predictions and criterion measures in opposite halves, and averaging these correlations to get an average validity coefficient. In essence, using the data from one half to predict the behaviour of the other half.

Using these linear regressions models, it is intended to apply estimates of utility and disutility (i.e. the extent to which a particular selection outcome gives a positive or negative return to the employer) in order to establish the optimal cut-off value of predictions to maximise the total expected utility. This cut-off value will be expressed as a percentage improvement on the total utility obtainable when the predictive equation is not used.

The design also includes an analysis of the effects of moderator variables (these are variables which interact with the predictors). One of the most important in this study is job complexity. Others which have been shown to be significant in other studies, and which will be addressed in this study, are race and sex.

D. RESEARCH HYPOTHESES

While this research in more empirical than theory based, it is possible to formulate at least two research hypotheses based on past findings.

1. Hypothesis 1

Age, educational level and mental ability/aptitude will be amongst the variables found to be significant predictors of criterion performances.

2. Hypothesis 2

Validities of the predictor/criterion regression equations will tend to increase as the level of complexity of the Navy rating increases.

IV. METHOD

This chapter describes the execution of the research. On a section by section basis it deals with: the ratings selected, the data base, the variables used, the subjects chosen and the statistical analyses employed.

A. NAVY RATINGS SELECTED

As has already been mentioned, the ratings in this study are Ships Serviceman, Personnelman and Aviation Technician. They were chosen because they represent varying degrees of job complexity. According to a complexity scale (see Sands, 1979) these ratings have values of :

a. Ships Serviceman : 40

b. Personnelman : 67

c. Aviation Technician : 95

The whole scale has a range from 10 to 99, with a median value of 70, and the most complex rating assigned a 99. These ratings therefore are representative of the spread of complexity of ratings in the US Navy.

Admittedly, the ratings differ in terms of the nature of the work as well as in complexity. This could well be a significant cause of rating differences and may influence the outcome of the study. A more ideal approach would have been to select different levels of complexity within the same type of employment. However, the data did not allow such a fine classification of personnel.

Another variable which will obviously have an effect on the outcome of this study is the fact that for all ratings an existing selection procedure is in use. It requires that entrants exceed particular ASVAB subtest scores, as well as reaching a certain SCREEN score. The SCREEN score is explained in detail later, but it is a probability which represents the chances of the applicant completing his period of service (Sands, 1979). The probability is derived from a consideration of race, AFQT, marital status and high school diploma status. Such prescreening will tend to restrict the range of both predictor and criterion variables which will have an effect on validity coefficients.

A general description of each rating, which is taken principally from the US Navy's "Enlisted Career Guide 1981-82", follows.

1. Ships Serviceman (SH)

At a primary level, depending upon their specialty, the SH may work in a barber shop, ship's store, laundry, dry cleaning plant or office. Subsequently they may assume more responsibility in supervisory and managerial roles associated with retail stores, commissary stores and laundry/dry cleaning plants. Among the distinguishing characteristics of this particular rating is an emphasis on basic interpersonal skills. While the actual employment may vary in physical surroundings and requirements, the jobs all have a lot to do with dealing and interacting with other people. An affinity for dealing with people is therefore essential. Although there is not a requirement for high level cognitive skills, it is nonetheless important for SHs to be of above average ability in arithmetic and in such things as record-keeping and detail work. After recruit training, the SH attends a 6 weeks' course in the trade. There are about 5,000 personnel employed in this rating at present.

2. Personnelman (PN)

The Navy's Personnelman is a rating which specialises in clerical and counselling duties with respect to personnel. They provide information and counselling to enlisted sailors related to Navy occupations, opportunities for general education and job training, promotion requirements, and in rights and benefits. After recruit training, the PN completes a 6 to 7 week technical school. There are about 7,000 personnel employed in this rating at present.

3. Aviation Technician (AT)

The full title of this rating is Aviation Electronics Technician. As the title suggests, sailors in the rating are employed maintaining advanced radio, radar and electronics equipment that are carried on board aircraft. Within this broad area of responsibility, there are three categories of employment. The AT may be involved in testing and analysis of equipment, its maintaince and repair, and in related administrative tasks. After recruit training, the AT attends up to 6 months of full-time schooling, and depending upon whether the AT is a four or six year obligator, he/she goes on to do a further 6 months of advanced first term avionics. The rating employes about 10,000 personnel at present, and there are shortages of ATs.

B. DATA BASE

The cohort data base was assembled by the Defense Manpower Data Centre (DMDC) for the Administrative Science Department of the Navy Postgraduate School (NPS). Cohort data on some 206,000 non-prior service sailors were obtained through the merging of three usually separate Navy files. These files are referred to as the DMDC file, the Advancement file, and the Navy Health and Research Center (NHRC) file.

The DMDC file contains pre-enlistment variables as well as entry dates and current status. It was assembled by linking various copies of the Enlisted Master Record across time using Social Security number. This number was deleted in the final cohort file in the interests of privacy.

The Advancement file records variables related to the promotion of individuals through the advancement system. For example, this file provided values for the Final Multiple, a composite score, which determines whether the sailor is promoted.

The NHRC file contains many of the same pre-entry variables as the DMDC file, but records additional pre- and post-entry variables. Amongst these are aptitude scores on the Armed Services Vocational Aptitude Battery (ASVAB) subtests and percentile scores on the Armed Forces Qualification Test (AFQT). The file also contains such things as the total number of absenses without leave (AWOLs), demotions, desertions and the times for promotion to various grades (eg number of days to E4). NHRC is the largest of the three files.

The cohort file included data from all enlisted entrants to the Navy between 1 September 1976 and 31 December 1978. No special data collect or monitoring of these subjects was set up; rather snapshots of their progress were taken in September 1982 from the various files to create the NPS cohort file. The time period allowed each enlistee the oppurtunity of serving about four years, with some being able to serve as long as six years.

C. VARIABLES IN THE DATA BASE

There are over 260 individual variables in the cohort data base, and those relevant for this study are described below.

1. Criterion Variables

Three criterion variables are used. One is unitary while the others are composites. The singular criterion is total length of service in months. This is available on all members of the cohort and is a true metric scale suitable for regression analysis. The second criterion is a positively oriented ordinal scale. Three mutual exclusive and collectively exhaustive categories of criterion outcomes are identified. Those individuals who did not complete four years of service for other than medical, death or officer entrance reasons, were grouped in the first category. Those who completed four years' service were placed in the second category. Those who served four years and were rated (i.e. became qualified in the rating), graded E4 and recommended for re-enlistment were placed in the third category. For the purposes of statistical analyses, these three categories were given the values "10", "20" and "30" respectively. Since this scale attempts to differentiate the better sailors from the others, it is known as the 'goodguy' scale. The third criterion is negatively oriented and is known as the 'badguy' scale. The first category includes those who did not complete four years of service, who had been discharged for negatively oriented reasons, and had either desertion, AWOL or confinements recorded against them and received a value of "10". The second ordinal category includes those who did not complete four years of service, or those who had demotions or were not recommended for re-enlistment regardless of their length of service and was valued "20". The remainder of the group, who did not exhibit any particularly negative behaviours, were placed in the third category which was valued "30".

While these criteria could be defined in different ways, they seem to be consistent with the three requirements suggested in Chapter II.

2. Predictor Variables

The following potential predictor variables are contained in the cohort file. They are:

a. age at entry;

b. marital status;

c. highest education level reached;

d. number of dependents;

e. various ASVAB subtest scores (in raw score form);

f. AFQT percentile scores;

g. groupings based on AFQT scores;

h. entry pay grade; and

i. SCREEN score (which is the probability of completing the period of enlistment based on education, AFQT and race).

These variables have all been previously associated with predicting criterion measures in military and non military settings. They are consistent with the three requirements for predictors mentioned in the literature survey.

3. Moderator Variables

As was pointed out earlier, it is suspected that complexity plays a moderating role. This has been accounted for, partially at least, by choosing three ratings which cover a wide span of complexity, and for which different selection procedures will be developed.

Race is another variable which may play a moderating role. Provided there are no adverse impacts for minorities, as measured by hiring rates, it may be possible to apply the same selection procedures across race, to the rating as a whole. Alternatively, selection procedures based specifically on race groupings may produce a better overall result. In any case, previous research emphasises wisdom of taking race into account.

Sex is another variable which is believed to play a moderating role in this research. In this study, there is a sufficient number of white females to perform analyses on two rating groups (these are PN and AT).

4. Screening of the Variables

While the accuracy of data entries into the original three data files is unknown, it is assumed to be very high. However, some cases were identified in the cohort data set which had impossible values. For example, no entrant to the Navy can be less than 17 years old, but some cohort members were listed with an entry age less than 17. These cases were excluded. Similarly, each of the ASVAB subtests has a known maximum number of items, thus cases with scores higher than these limits were excluded. Because the criterion measures all required that each sailor have the opportunity of serving for at least four years, then those sailors whose enlistment dates precluded serving this period were excluded.

The exclusion of cases for these reasons was 'listwise'. That is, the case and all its variable values was deleted. The percentage of cases deleted in this way amounted to less than 5%.

5. Estimates of Utility

As was argued in Chapter II, if a significant relationship between predictors and the criterion is found, then estimates of utility for various selection outcomes become important for deciding the optimum cut-off on the predictor to achieve the best overall selection result. This involves determining the cut-off level on the predictor for which total utility is maximised (see Table II and related discussion). This section discusses the derivation of the estimates of utility.

To estimate the utility associated with each of the four possible selection outcomes it is assumed that the Billet Cost Model (see Griffin, 1981, for an explanation of the cost components of this model) provides reasonable estimates. As is discussed by Campbell (1976), it is the inter-cell ratios which are important rather than the absolute values themselves. The Billet Cost Model provides a method of approximating these ratios.

For correctly predicting a successful sailor, who will complete four years of service, the total marginal cost is taken as the best estimate of utility. This is based on the assumption that the Navy compensates sailors at the full value of their marginal product. This is probably a conservative estimate as a recent paper (Butler,1982) has suggested. Butler has calculated for the Electronic Technician Rating, based on the amount of training dollars the Navy spends on these sailors and their expected total service, that the Navy must expect a return in product considerably in excess of rating billet cost. Nevertheless, the total billet cost figure is accepted as the best estimate of utility for this outcome (i.e. correctly predicting success).

An estimation of the disutility of predicting success, when in fact the sailor fails the criterion, is provided by the rating cost estimates of Balis and Clay-Mendez (1982). Basically these data come from the Billet Cost Model, but include only a few of the elements of the original model. The Balis and Clay-Mendez cost amounts are the average replacement costs by ratings across so called quality gradings. The cost of the replacement of the highest quality grading is used in the present study as the estimate of the disutility of falsely predicting success.

Relevant product/cost figures are not available for the utilities associated with applicants who are rejected by the prediction device. However, these utilities are derived from the above estimates. In the case of the applicant correctly rejected (he/she would have failed on the criterion measure), it is argued that this selection outcome has the same utility as the disutility associated with the entrant who fails the criterion. Therefore, the 'Balis' cost data is used as the utility estimate for this selection outcome.

The disutility of the rejection of the potentially successful applicant is derived as follows. The disutility associated with this outcome depends on the recruiting market and whether it is easy or difficult to get sailors. In some circumstances it could be that there are so many good quality applicants that rejecting potentially successful applicants has little or no disutility. One extreme value for this utility, therefore, could be zero. At the other extreme, the disutility of rejecting these applicants could be equal in magnitude to the utility of enlisting a successful applicant. The estimate of the disutility of rejecting a potentially successful aplilicant is taken as the average of these two extremes. The utility values, by rating, are given in Table IV .

TABLE IV

Relative Utilities by Rating

RATING	PREDICTOR *** SUCCESS *** Criterion Success	Criterion Failure	PREDICTOR *** FAILURE *** Criterion Success	Criterion Failure
SH	19260	-13077	-9630	13077
EN	18488	-14495	-9244	14495
AT	22297	-21210	-11149	21210

The above utilities are used to derive cut-off scores for all three criteria. For the Length of Service (LCS) criterion they are used directly as the utilities for each of the four possible selection outcomes. On the goodguy and badguy scales they are used as the basis of developing utilities as described below.

TABLE V

Relative Goodguy Utilities by Rating

	PREDICTOR *** SUCCESS *** Criterion 30	20	10	PREDICTOR *** FAILURE *** Criterion 30	20	10
SH	28890	19260	-13077	-19260	-9630	13077
EN	27732	18488	-14495	-18488	-9244	14495
AT	33446	22297	-21210	-22297	-11149	21210

For the goodguy scale, there are three criterion outcomes, which for the purposes of establishing a cut-off, gives a total of six possible criterion/predictor combinations. The scale values 20 and 10 are taken to have utilities as shown in the above table. The utility of correctly predicting a goodguy 30 is computed to be one and a half times the utility of correctly predicting a goodguy 20. As was done in developing the above values, the disutility of rejecting a sailor who would have been a goodguy 30 is taken as the simple average of its extreme values. As can be seen in Table V, the figure for SH is the average of 28890 and 9630.

TABLE VI

Relative Badguy Utilities by Rating

	PREDICTOR *** SUCCESS ***			PREDICTOR *** FAILURE ***		
	Criterion			Criterion		
	30	20	10	30	20	10
SH	19260	13077	-19616	-9630	-13077	16347
FN	18488	14495	-21743	-9244	-14495	18119
AT	22297	21210	-31815	-11149	-21210	26513

Similarly, the badguy utilities, shown in Table VI, were derived starting from those in Table IV. These four utilities represent the outcomes for bad guy criterion scores of 20 and 30. Again, it was figured that the utility of correctly identifying and rejecting badguy 10s was one and a half times the utility of correctly identifying a 20. As for the goodguy scale a simple average gave the utility for the other outcome.

D. SUBJECTS CHOSEN

Within the NPS cohort file there are three possible indicators of a sailor's rating. These are as follows. Firstly, there is a rating recorded at the time of entry. Secondly, when the sailor attempts the rating examination, the resulting rating is recorded. Thirdly, the DMDC file contains a rating which is the sailor's official designation in terms of rating. These rating variables are known, respectively, as (1) entry rate, (2) examination rate and (3) DMDC rate.

It would be an ideal situation if all three rating variables agreed. However, there is considerable variation in the recordings of rating across these variables. In fact there are a total of seven different combinations when it comes to selecting suitable subjects for a particular rating. These seven possible combinations are shown in the Table VII.

TABLE VII

Possibilities for Selecting Subjects within a Rating

	Rating Indicators		
Category	Entry	Exam	DMDC
1.	Yes	Yes	Yes
2.	Yes	Yes	No
3.	Yes	No	Yes
4.	Yes	No	No
5.	No	No	Yes
6.	No	Yes	Yes
7.	No	No	Yes

Note 1: 'Yes' means that the data file indicated that the sailor was a member of the rating according to the particular rating indicator.

Note 2: In choosing subjects who were representative of rating membership, all the above Categories except # 4. were used as selection criteria.

There are a number of ways for determining which subjects in the listed categories are representative of a rating. One method could be to select only those subjects who were coded in the rating for all three rating selection variables. But this approach would seriously restrict the sample size and ,indeed, it would probably not produce a repesentative sample of the members of the rating.

Of the seven possible categories listed in Table VII, it was decided to initially select on all seven, then to exclude those sailors in category 4. That is, the exclusions were those Navy entrants for whom the only indication of membership of the rating was their entry code. It was reasoned that while these sailors showed an interest in the rating at some stage prior to entry, there is no indication that he/she had ever had any work experience in the rating. From a concurrent validity point of view, it is necessary for the subject sailors selected for a rating to at least experience it and to have their rating performances influenced by that experience. Presumably the sailors who were coded in the rating through the examination rate or through a DMDC code have had sufficient experience in the rating to be considered representative of the rating.

After these variable screens were applied, and the rating categories identified, subjects were chosen within each rating. These selections are shown by race and sex in Table VIII. As will be seen in the results section, it was decided that the numbers in some of the race/sex combinations were too small to perform realistic analysis. There are sufficient numbers of men in each race grouping. However there are enough white women in only the PN or AT ratings to perform regression and other analyses. Therefore, in all subsequent analyses, three groupings for SH, and four each, for PN and AT, are reported.

TABLE VIII

Rating Selectees by Rating, Race and Sex

RATING	Sex	White	Black	Others	TOTAL
SH	Male	1330	572	169	2071
SH	Female	28	7	2	37
EN	Male	1263	288	112	1663
EN	Female	479	70	15	564
AT	Male	3400	176	114	3690
AT	Female	243	14	3	260

Note 1: The number of personnel are shown in each cell.

E. STATISTICAL PROCEDURES

1. General

Now that the criterion, predictor and moderator variables have been indicated and the subjects identified, the statistical procedures will be described. The Statistical Analysis System (SAS) package (SAS, 1979) was used for the analyses, along with some FORTRAN programs written by the author. Sample programs are listed in Appendix A.

2. Descriptive Analyses

These are given by moderator variables, and consist of the score and/or mean values on predictor and criterion variables. No statistical tests are applied to these values, since where significant differences occur these are "captured" in the subsequent analyses.

3. Predictive Analyses

For each rating group the following steps are applied:

a. stepwise regression is used to identify significant predictors;

b. these significant predictors are used in 'ordinary least squares' regression to develop 'double cross-validation' validity coefficients; and

c. regression weights are used to score each case on the respective predictor equations.

When the group size was small, the number of predictors used in the subsequent stepwise analysis, was limited to one variable to every 20 personnel. As a consequence, the 'other' racial category for ATs was limited to the first five variables selected in the stepwise procedure.

4. Estimation of Cut-offs

The 'ordinary least squares' regression equations are used to estimate optimum cut-offs as is described below.

In order to use the utility values to estimate the most ideal cut-offs on the predictors, it is first necessary to consider the LOS criterion measure as a dichotomous variable. Thus the sample is divided into two groups: those with less than four years and those with four years or more of service. The two three-point criterion scales were not recoded.

As is illustrated in the Appendix A, during the double cross-validation SAS run on the computer, criterion measures predictor scores and the variable race were output to a separate computer file which could be accessed by the FORTRAN program. The predictor score output at this stage was derived in the following way: the separate regression coefficients, in each of the cross-validating samples, were used to separately score all the cases producing two predicted scores for each case. A simple average was taken, which represents the predictor score output for later FORTRAN analysis.

The method for estimating the most appropriate cut-off score was explained in Chapter II. In summary, the cutting score chosen is the one which maximises equation 4.1.

$$Uc = \text{Sum of } (Pr(i) * Ut(i)) \qquad \text{(eqn 4.1)}$$

where:

Uc is the Total Utility for Cutting Score c.

Pr(i) is the Probability of Outcome (i).

Ut(i) is the Utility of Outcome (i).

The results of these analyses will be presented in graphical form. In order to get some frame of reference for evaluating the value of the chosen cut-off score, the cut-off is expressed as a percentage change from, what is known as base rate. The base rate is simply defined as the value of equation 4.1 when the cut-off is set so low that every applicant is accepted. Unfortunately, base rate is not the true utility of the current selection, since it is represents the assessment of the utility only for those who are selected by these procedures, ignoring applicants who are rejected by them. As is discussed in Chapter VI, the overall utility of a selection device must take into account the costs involved in setting up and maintaining the system.

Thus, base rate, as defined here, is a utility estimate which ignores these external costs, but it provides an index figure for judging each potential cutting score. This index for each score is expressed as a percentage change, either positive or negative, from base rate.

V. RESULTS

This chapter presents the thesis results and is organised into three sections. The first is descriptive, providing breakdowns on criterion, predictor and moderator variables. The second section details the predictive results. This includes the variables selected for multiple regression, and also the results of cross validation. In the last section, the results of the application of utilities to estimate appropriate cutting scores, are presented.

A. DESCRIPTIVE RESULTS

1. SH Groups

Shown in Table IX are the means and/or frequencies by category for potential predictor and criterion variables across the three male SH racial groupings. It is interesting to note that the length of service is greatest for non-white races. On the other hand, both the goodguy and badguy scales indicate that proportionally more of the white group have extreme values than do the other two races.

Among the predictor variables there are also some interesting trends. White personnel seem to be younger with lower educational level. However, with some exceptions, the white group's performances on the ASVAB subtests are better. This finding is consistent with this group's better AFQT scores.

2. EN Groups

The same basic trends noted for SH men are evident for EN men, as is shown in Table X. The white men are slightly younger, they exhibit 'good' rather than 'bad'

TABLE IX

Criterion and Predictor Values for SH Men

Variable		White	Black	Other
Number of Cases		1330	572	169
Mean LOS: months		44.99	49.31	50.82
Goodguy Scale	30	24%	16%	12%
	20	33%	49%	58%
	10	43%	35%	30%
Badguy Scale	30	23%	17%	14%
	20	51%	66%	72%
	10	26%	17%	14%
PREDICTOR MEANS				
Entry Age: years		18.7	19.4	21.0
Marital Status		1.3	1.4	1.5
Highest Education		11.6	11.8	12.2
Dependents (number)		.03	.05	.02
AFQT Percentile		46.6	34.7	32.3
AFQT Group		5.3	4.5	4.0
Entry Paygrade		1.1	1.2	1.1
SCREEN Score		80.2	79.4	80.2
ASVAB Subtests:				
Attention to Detail		14.6	13.8	12.9
Numerical Operations		32.1	28.1	27.3
Auto. Information		10.1	7.3	6.9
Attentiveness Scale		9.7	10.8	9.9
General Science		10.4	8.5	7.8
Electronics Scale		7.1	8.3	8.5
Math. Knowledge		11.4	9.3	8.9
Space Perception		12.1	11.1	11.0
Maintenance Scale		9.9	9.5	9.6
Electronic Info.		17.9	15.2	14.6
Arithmetic Reasoning		12.9	10.7	9.5
General Information		9.4	7.8	6.5
Word Knowledge		19.5	17.5	15.1
Shop Information		12.7	9.7	9.4
Combat Scale		15.9	14.1	13.4
Mech. Comprehension		9.5	7.4	6.9

Note 1: Goodguy: 30 .. served 4 years, promoted E4 and recommended for reenlistment.
20 .. served 4 years.
10 .. remainder after 20 and 30.

Note 2: Badguy: 30 .. remainder after 10 and 20.
20 .. minor negative indicators.
10 .. major negative indicators.

Note 3: Goodguy and Badguy percentages sum to 100%.

Note 4: AFQT groups: values 1 to 8, represent categories (in order) 5, 4C, 4B, 4A, 3B, 3A, 2 and 1.

Note 5: Marital Status: married (2); other (1).

TABLE X

Criterion and Predictor Values for PN Groups

Variable		MALE White	Black	Other	FEMALE White
Number of Cases		1263	288	112	479
Mean LOS: months		48.72	49.53	51.94	45.67
Goodguy Scale	30	17%	9%	12%	21%
	20	48%	52%	63%	36%
	10	35%	39%	25%	43%
Badguy Scale	30	17%	9%	14%	22%
	20	72%	81%	79%	74%
	10	11%	10%	8%	4%
PREDICTOR MEANS					
Entry Age: years		19.8	20.7	21.1	20.0
Marital Status		1.4	1.4	1.5	1.5
Highest Education		12.1	12.2	12.7	12.2
Dependents (number)		.07	.06	.08	.05
AFQT Percentile		63.0	49.3	34.8	63.9
AFQT Group		6.2	5.5	3.9	6.3
Entry Paygrade		1.4	1.5	1.1	1.4
SCREEN Score		84.3	82.7	82.3	None
ASVAB Subtests:					
Attention to Detail		15.0	14.0	11.8	16.1
Numerical Operations		35.4	30.6	25.6	37.8
Auto. Information		11.3	8.4	6.3	7.6
Attentiveness Scale		11.5	12.6	9.0	12.6
General Science		12.6	10.5	7.8	11.9
Electronics Scale		7.9	8.9	6.9	5.2
Math. Knowledge		13.7	11.1	8.8	13.8
Space Perception		12.6	11.1	9.4	12.9
Maintenance Scale		9.1	9.1	7.5	6.1
Electronic Info.		19.8	16.7	13.4	16.8
Arithmetic Reasoning		15.0	12.4	9.3	14.3
General Information		10.6	8.6	6.2	8.1
Word Knowledge		23.7	22.0	14.7	24.6
Shop Information		13.6	10.7	8.7	9.9
Combat Scale		15.8	14.6	11.7	14.2
Mech. Comprehension		11.0	8.1	6.7	9.4

Notes: As for Table IX.

behaviours, and are better performers on the ASVAB and the AFQT than are the men in the other racial groups.

The scores of the white PN women, not surprisingly, are closest to the scores of the white PN men. The women, however, have considerably shorter average lengths' of service, but exhibit a higher proportion of 'good' rather

TABLE XI

Criterion and Predictor Values for AT Groups

Variable		MALE White	MALE Black	MALE Other	FEMALE White
Number of Cases		3339	172	109	242
Mean LOS: months		53.32	53.39	51.16	45.37
Goodguy Scale	30	18%	10%	18%	14%
	20	52%	57%	48%	44%
	10	30%	33%	34%	42%
Badguy Scale	30	19%	10%	16%	15%
	20	76%	83%	82%	76%
	10	5%	7%	2%	3%
PREDICTOR MEANS					
Entry Age: years		19.1	20.5	20.2	20.6
Marital Status		1.3	1.4	1.5	1.5
Highest Education		12.0	12.2	12.2	12.5
Dependents (number)		.06	.09	.08	.08
AFQT Percentile		72.5	56.5	55.6	79.7
AFQT Group		6.7	5.9	5.6	7.0
Entry Paygrade		1.9	1.8	1.5	1.3
SCREEN Score		86.2	83.6	83.9	None
ASVAB Subtests:					
Attention to Detail		15.2	14.5	15.5	16.7
Numerical Operations		35.5	31.4	33.5	40.5
Auto. Information		14.7	10.9	11.1	10.0
Attentiveness Scale		9.4	10.9	9.3	12.4
General Science		14.9	12.7	12.5	14.8
Electronics Scale		11.0	11.5	10.5	9.1
Math. Knowledge		16.0	13.9	14.5	16.6
Space Perception		14.9	13.2	14.0	15.3
Maintenance Scale		12.0	10.7	10.0	8.7
Electronic Info.		24.2	22.0	21.0	20.3
Arithmetic Reasoning		16.3	14.0	13.8	16.7
General Information		11.6	10.1	9.1	9.6
Word Knowledge		24.2	21.5	19.6	26.6
Shop Information		16.2	12.9	13.2	12.1
Combat Scale		16.4	14.6	14.4	15.8
Mech. Comprehension		14.5	10.7	11.7	11.5

Notes: As for Table IX.

'bad' behaviours. On ASVAB subtests the women, when compared with white men, are better on some scales (e.g. Numerical Operations and Word Knowledge), poorer on some others (e.g. Auto Information and Electronics Scale) and about the same on the remainder. On educational level, the women have entry educational levels which are within the extremes of the three male racial categories on this variable. Incidentally, it is the male 'other' race category which has the highest educational level for the PN rating.

3. AT Groups

For the AT rating perhaps the most striking fact is the proportion of ATs who are white (see Table XI, about 92%). This is considerably larger than the equivalent percentages for PN (76% white), and SH (64% white).

The trends in the AT data are consistent with those already mentioned for PNs and SHs. Again, the white group is younger, with lower education level and generally better AFQT scores. However, for AT there are proportionally more Blacks in the negative category of the 'bad guy' scale (i.e. those scores of '10'). Women again have superior performances on some ASVAB scales and, for this rating, they have the highest education level.

B. PREDICTIVE RESULTS

In this section the results of the stepwise regressions, in the form of the variables selected and the signs of the respective coefficients, are presented. These are followed immediately by validity estimates for the predictive models which are constructed. Regression coefficients and statistics are not listed in the body of the thesis: however, interested readers are directed to Appendix B where they are given in full.

1. Length of Service Criterion

For the SH rating, the variables selected as predictors on the length of service criterion by race, according to the sign of the coefficient, are shown on Table XII. Across races all the signs are consistent for predictors

TABLE XII

Stepwise Regression for SH on LOS

Variable	White	Black	Other
Entry Age: years	n	n	Pos
Marital Status	Pos	Pos	n
Highest Education	n	Pos	n
Attention to Detail	n	n	Neg
Auto. Information	Neg	n	n
Attentiveness Scale	Neg	Neg	n
General Science	n	n	Pos
Arithmetic Reasoning	n	n	Neg
General Information	Neg	n	Neg
Mech. Comprehension	n	n	Neg
Combat Scale	n	Pos	n
AFQT Group	Neg	Neg	n
Entry Paygrade	Pos	n	n
SCREEN Score	Pos	Pos	n
Dependents (number)	Neg	n	n

Note

Pos means the coefficient was positive.
Neg means the coefficient was negative.
'n' means the variable was not selected.

which entered the equation in more than one racial grouping. Entry age, marital status, educational level and entry paygrade are all positively related to length of service. Longer service is associated with higher values on these variables. Nearly all the predictive ASVAB subtests, and the AFQT groupings, are negatively related to length of service.

As shown in Table XIII, there are some differences in the signs associated with predictors for the PN groups for the length of service criterion. For whites and blacks, higher education is associated with shorter service, while for the 'other' race category, the reverse is true. There are mixed signs for the ASVAB subtests, but the SCREEN score consistently has a positive sign. Some ASVAB predictors for length of service, in the female group, have negative signs. Perhaps this is not surprising, because as was noted earlier, women tended to have shorter service but better ASVAB subtest scores than did men.

TABLE XIII

Stepwise Regression for PN on LOS

Variable	MALE White	MALE Black	MALE Other	FEMALE White
Entry Age: years	Pos	n	n	n
Marital Status	Pos	Pos	n	n
Highest Education	Neg	Neg	Pos	Neg
Attention to Detail	n	n	n	Neg
Numerical Operations	n	Neg	n	n
Electronics Scale	Pos	n	n	Pos
Electronic Info.	Pos	n	n	n
Arithmetic Reasoning	Neg	n	n	Neg
Word Knowledge	Neg	n	n	Neg
Combat Scale	n	Neg	n	n
Shop Information	Pos	n	n	n
AFQT Percentile	n	Neg	n	n
Entry Paygrade	n	n	n	Pos
SCREEN Score	Pos	Pos	n	n
Dependents (number)	Pos	Pos	Neg	n

Note
Pos means the coefficient was positive.
Neg means the coefficient was negative.
'n' means the variable was not selected.

TABLE XIV

Stepwise Regression for AT on LOS

Variable	MALE White	MALE Black	MALE Other	FEMALE White
Entry Age: years	n	n	Pos	n
Marital Status	Pos	Pos	n	Neg
Highest Education	Neg	Neg	n	n
Numerical Operations	Neg	n	n	n
Auto. Information	n	n	Neg	n
Attentiveness Scale	n	Pos	n	n
General Science	Neg	Neg	n	n
Electronics Scale	Pos	n	n	Neg
Math. Knowledge	Neg	n	n	n
Space Perception	Neg	n	n	n
Maintenance Scale	n	n	Pos	Pos
Electronic Info.	n	n	Pos	Pos
Shop Information	n	Neg	n	n
AFQT Percentile	Pos	n	n	n
AFQT Group	Neg	n	n	n
Entry Paygrade	Pos	n	n	n
SCREEN Score	Pos	n	Pos	n

Note
Pos means the coefficient was positive.
Neg means the coefficient was negative.
'n' means the variable was not selected.

Like the SHs, there is a consistency of signs for stepwise entered predictors for the AT rating (see Table XIV) for length of service. However, the direction of the signs, for some variables, is different. For example, educational level for two racial groupings, is now negatively related to length of service. This sign change is carried over into the ASVAB subtests. Being such a technically complex rating, one would expect Science and Mathematics to be positively related to length of service. However both of these scales have negative signs. Consistent with some of the other stepwise regressions, SCREEN score is positively related to length of service.

2. <u>Goodguy Criterion</u>

Table XV shows the variables selected by race on the goodguy scale for the SH rating. This is the smallest

TABLE XV

Stepwise Regression for SH on Goodguy

Variable	White	Black	Other
Marital Status	Pos	Pos	n
Highest Education	Pos	n	n
Auto. Information	n	n	Neg
Electronics Scale	Neg	n	n
Math. Knowledge	Pos	n	n
Electronic Info.	n	n	Neg
Combat Scale	n	Pos	Pos
AFQT Percentile	n	n	Neg
AFQT Group	n	Neg	n
SCREEN Score	n	Pos	Pos

<u>Note</u>

Pos means the coefficient was positive.
Neg means the coefficient was negative.
'n' means the variable was not selected.

number of variables selected for any of the stepwise regressions. The signs are all consistent with the predictor selections for SH on length of service. However, it is interesting to note that AFQT percentile, for the 'other' racial category, has a negative sign: lower AFQT percentiles are predictive of better goodguy performances.

TABLE XVI

Stepwise Regression for PN on Goodguy

Variable	MALE White	MALE Black	MALE Other	FEMALE White
Marital Status	Pos	n	n	n
Highest Education	Neg	Neg	Pos	n
Attention to Detail	n	n	n	Neg
Numerical Operations	n	Neg	n	n
Electronics Scale	Pos	Neg	n	n
Math. Knowledge	n	Pos	n	n
Mech. Comprehension	Pos	n	n	n
Electronic Info.	n	Pos	n	n
Arithmetic Reasoning	Neg	n	Pos	Neg
Word Knowledge	Neg	n	n	Neg
Combat Scale	n	n	Pos	n
General Information	n	Pos	n	n
Shop Information	n	Neg	Neg	n
AFQT Percentile	n	n	Neg	n
SCREEN Score	Pos	n	n	n
Dependents (number)	Pos	n	Neg	n

Note

Pos means the coefficient was positive.
Neg means the coefficient was negative.
'n' means the variable was not selected.

As was the case on the length of service criterion for PNs, the level of education is negatively related to the goodguy scale. This is shown in Table XVI. ASVAB subtests selected have varying signs and for two racial groupings for men, the variable 'number of dependents' has entered the regression equation. The SCREEN score for white males is positively related to desirable goodguy behaviors.

For the AT groupings, on the goodguy scale, see Table XVII, there is little consistency regarding which variables are selected. Only three predictors were chosen in more than one grouping, and when this happened the signs are different across groups. It is interesting that for the largest group, white males, not a single ASVAB subtest is selected as a predictor.

TABLE XVII

Stepwise Regression for AT on Goodguy

Variable	MALE White	MALE Black	MALE Other	FEMALE White
Marital Status	Pos	n	n	n
Highest Education	Pos	Neg	n	n
Attention to Detail	n	n	Pos	n
Attentiveness Scale	n	Pos	n	n
General Science	n	Neg	n	n
Electronics Scale	n	n	n	Neg
Maintenance Scale	n	n	n	Pos
Electronic Info.	n	n	Neg	Pos
Word Knowledge	n	n	Pos	n
AFQT Percentile	n	Pos	Neg	n
AFQT Group	n	Neg	n	n
Entry Paygrade	Neg	n	n	n
SCREEN Score	n	n	Pos	n

Note
Pos means the coefficient was positive.
Neg means the coefficient was negative.
'n' means the variable was not selected.

3. Badguy Criterion

The SH rating, see Table XVIII, on this scale, shows a consistent trend for education to be positively related

TABLE XVIII

Stepwise Regression for SH on Badguy

Variable	White	Black	Other
Entry Age: years	n	Neg	n
Marital Status	Pos	n	n
Highest Education	Pos	Pos	Pos
Numerical Operations	Pos	n	n
Auto. Information	n	n	Neg
Electronics Scale	Neg	n	n
Arithmetic Reasoning	n	Neg	n
Combat Scale	n	Pos	Pos
Word Knowledge	Neg	n	n
AFQT Percentile	n	n	Neg
AFQT Group	n	n	Pos
Entry Paygrade	Pos	n	n
SCREEN Score	Pos	n	n

Note
Pos means the coefficient was positive.
Neg means the coefficient was negative.
'n' means the variable was not selected.

for all three racial groupings. Since the goodguy and the badguy scales are both scored in the same direction (i.e. the 'good' and 'bad' ends have the same numerical values), this means that higher education is associated with 'non-bad' behaviours. Several ASVAB subtests were selected, with differing signs. Entry paygrade and SCREEN score are both positive.

TABLE XIX

Stepwise Regression for PN on Badguy

Variable	MALE White	MALE Black	MALE Other	FEMALE White
Entry Age: years	n	Neg	n	n
Highest Education	n	n	Pos	Pos
Attention to Detail	n	n	n	Neg
Auto. Information	n	Pos	n	n
Space Perception	n	n	n	Pos
Electronic Info.	Pos	n	n	n
Arithmetic Reasoning	n	n	n	Neg
Shop Information	n	Neg	Neg	n
AFQT Percentile	n	n	Neg	n
AFQT Group	Neg	n	n	n
Entry Paygrade	n	Pos	n	n
SCREEN Score	Pos	n	n	n
Dependents (number)	Pos	n	Neg	Neg

Note
Pos means the coefficient was positive.
Neg means the coefficient was negative.
'n' means the variable was not selected.

For PNs, see Table XIX, entry age entered the regression only for blacks, and it has a negative coefficient. Education level is selected for two groups, the 'other' male race category and white women, and for both, the sign is positive: higher education level implies fewer undesirable behaviours. There are a variety of ASVAB subtests selected with differing signs, and it is interesting to note that for three of the four groups the number of dependents entered the regression. For white males, having more dependents suggest fewer negative behaviours, while for the 'other' races and white females, having more dependents is associated with more negative behaviours.

TABLE XX

Stepwise Regression for AT on Badguy

Variable	MALE White	MALE Black	MALE Other	FEMALE White
Entry Age: years	Pos	n	n	Pos
Marital Status	Neg	n	n	Pos
Highest Education	Pos	n	n	n
Auto. Information	n	Pos	Neg	n
Electronics Scale	n	n	n	Neg
Math. Knowledge	n	Neg	Pos	n
Electronic Info.	n	n	Neg	n
Arithmetic Reasoning	n	Pos	n	n
Combat Scale	n	n	n	Pos
AFQT Percentile	n	Pos	n	n
AFQT Group	Neg	Neg	n	n
Entry Paygrade	Neg	Neg	n	Neg
SCREEN Score	Pos	n	Pos	n
Dependents (number)	n	n	n	Neg

Note
Pos means the coefficient was positive.
Neg means the coefficient was negative.
'n' means the variable was not selected.

The final stepwise table, Table XX, shows the selected predictors for AT groupings on the badguy scale. The most consistent subtest finding across groupings is the entry paygrade, which has a negative sign. This variable is selected in three groupings and implies that the higher the entry pay grade the more likely are negative type behaviours.

4. Validity Estimates

Shown in the next series of tables are double cross-validition coefficients by rating, grouping and randomly selected sample. They are based on forming predictor equations from the stepwise procedures for each individual group. Validities for the "men all" category were the result of forming regression equations on all predictors which had been independently selected by male race groupings. The average validities given were calculated using Fisher's transformations for determining average weigthed correlation coefficients (see McNemar, 1963, pp. 139-140).

a. Length of Service

The length of service validities by individual sample are shown in Table XXI. All of them are significant except for the 'other' male race grouping, in both the PN and AT ratings, and the women in the AT rating. Most of the validities are greater than .20, but only one (AT blacks in Sample 1) is in excess of .40.

TABLE XXI

Length of Service Validities by Rating and Sample

RATING		Sample 1 n	Sample 1 r	Sample 1 p <	Sample 2 n	Sample 2 r	Sample 2 p <
SH							
MEN	All	1016	.2748	.0001	1022	.2806	.0001
	White	671	.1893	.0001	635	.2330	.0001
	Black	262	.3375	.0001	236	.2581	.0001
	Other	91	.3053	.0033	78	.3640	.0011
PN							
MEN	All	817	.2731	.0001	783	.2785	.0001
	White	606	.2881	.0001	620	.2894	.0001
	Black	158	.3519	.0001	125	.2963	.0008
	Other	66	.0546	.6634	46	.1294	.3919
WOMEN	White	241	.2060	.0013	238	.2939	.0001
AT							
MEN	All	1824	.2530	.0001	1799	.2803	.0001
	White	1676	.2617	.0001	1678	.2856	.0001
	Black	91	.4227	.0001	85	.3338	.0018
	Other	58	.2393	.0705	52	.2125	.1304
WOMEN	White	124	.1219	.1771	116	.2399	.0095

b. Goodguy Criterion

For each rating as shown in Table XXII, some of the validities by groups are statistically insignificant. The overall magnitude of these validities, compared to those for the length of service criterion, are smaller. The maximum validity is .378, while a number of them are less then .10.

TABLE XXII

Goodguy Validities by Rating and Sample

RATING		Sample 1 n	Sample 1 r	Sample 1 p <	Sample 2 n	Sample 2 r	Sample 2 p <
SH							
MEN	All	1016	.1052	.0008	1004	.1373	.0001
	White	681	.1586	.0001	649	.2176	.0001
	Black	262	.1268	.0403	296	.1593	.0060
	Other	83	.3673	.0006	73	.3916	.0006
PN							
MEN	All	817	.0914	.0090	783	.0666	.0627
	White	606	.1695	.0001	620	.1464	.0003
	Black	160	.1579	.0462	128	.1179	.1852
	Other	66	.0987	.4306	46	.2631	.0773
WOMEN	White	241	.1752	.0064	238	.1799	.0054
AT							
MEN	All	1824	.0421	.0723	1799	.0339	.1503
	White	1712	.0895	.0002	1688	.0303	.2128
	Black	91	.3046	.0033	85	.3503	.0010
	Other	58	.2577	.0508	52	.3464	.0119
WOMEN	White	125	.0004	.9969	118	.0514	.5807

TABLE XXIII

Badguy Validities by Rating and Sample

RATING		Sample 1 n	Sample 1 r	Sample 1 p <	Sample 2 n	Sample 2 r	Sample 2 p <
SH							
MEN	All	1016	.1713	.0001	1004	.2062	.0001
	White	671	.1967	.0001	635	.2589	.0001
	Black	270	.1178	.0531	302	.1890	.0010
	Other	91	.3652	.0004	78	.3254	.0037
PN							
MEN	All	817	.0963	.0059	783	.0980	.0061
	White	606	.1452	.0003	620	.1174	.0034
	Black	160	.2852	.0003	128	.2049	.0204
	Other	66	.1102	.3784	46	.1717	.2538
WOMEN	White	241	.1815	.0047	238	.1896	.0033
AT							
MEN	All	1824	.2293	.0001	1799	.1999	.0001
	White	1676	.2513	.0001	1664	.2086	.0001
	Black	91	.2741	.0086	85	.3670	.0006
	Other	58	.3728	.0040	52	.4464	.0009
WOMEN	White	125	.2688	.0024	118	.2400	.0089

c. Badguy Criterion

For this criterion, validities seem to be higher than for the good guy criterion, as shown in Table XXIII. Only four of the 28 sample validities are insignificant at the .01 level of significance. The highest validity of any presented so far (.4464) occurs for one of the 'other' race samples (for ATs).

5. Average Validities

Using the Fisher method, average validities were calculated from the validities just presented and these are shown in the next two tables. Table XXIV, shows the average

TABLE XXIV

Average Validities by Group and Criterion

RATING		CRITERION ICS	Goodguy	Badguy
SH				
MEN	All	.2777	.1212	.1879
	White	.2107	.1876	.2272
	Black	.2959	.1441	.1556
	Other	.3327	.3787	.3470
EN				
MEN	All	.2757	.0793	.0971
	White	.2888	.1578	.1312
	Black	.3277	.1412	.2500
	Other	.0852	.1669	.1354
WOMEN	White	.2502	.1775	.1855
AT				
MEN	All	.2666	.1380	.2148
	White	.2737	.0602	.2301
	Black	.3801	.3268	.3197
	Other	.2262	.3002	.4082
WOMEN	White	.1796	.0252	.2549

validity across each of the samples by racial groupings and ratings. The largest average validity occurs for the AT rating (.41) in the 'other' racial group on the bad guy criterion. The second highest validity is also for an AT rating (.38); it is for the black racial group.

Table XXV, lists grand average validities for men, for rating by criterion. For length of service, all three rating average validities are in excess of .25, with a slightly increasing trend in size of coefficent as the

TABLE XXV

Validities for Males by Rating and Criterion

	CRITERION		
RATING	ICS	Goodguy	Badguy
SH	.2515	.2046	.2277
PN	.2707	.1562	.1517
AT	.2762	.0867	.2450

rating complexity increases. For the goodguy scale, the rank order of validities by magnitude is reversed: validities decrease with rating complexity. The average validity for ATs on this criterion is particularly small. There is no consistent ranking by complexity for the other criterion, badguy. The PNs validity remains about the same as for the goodguy scale, as does the SH validity. The average validity for the ATs on this scale has increased considerably, up to .245.

C. ESTIMATING CUT-OFFS

The method for estimating the cut-offs was described in the previous chapter. However, to illustrate the method, the calculation of the first few values in Figure 5.1 is given here. It should be noted that most of the figures in this section have two or more graphs drawn on them. This is done in the interests of economy rather than to imply they are related: each graph is based on a separate predictor criterion relationship.

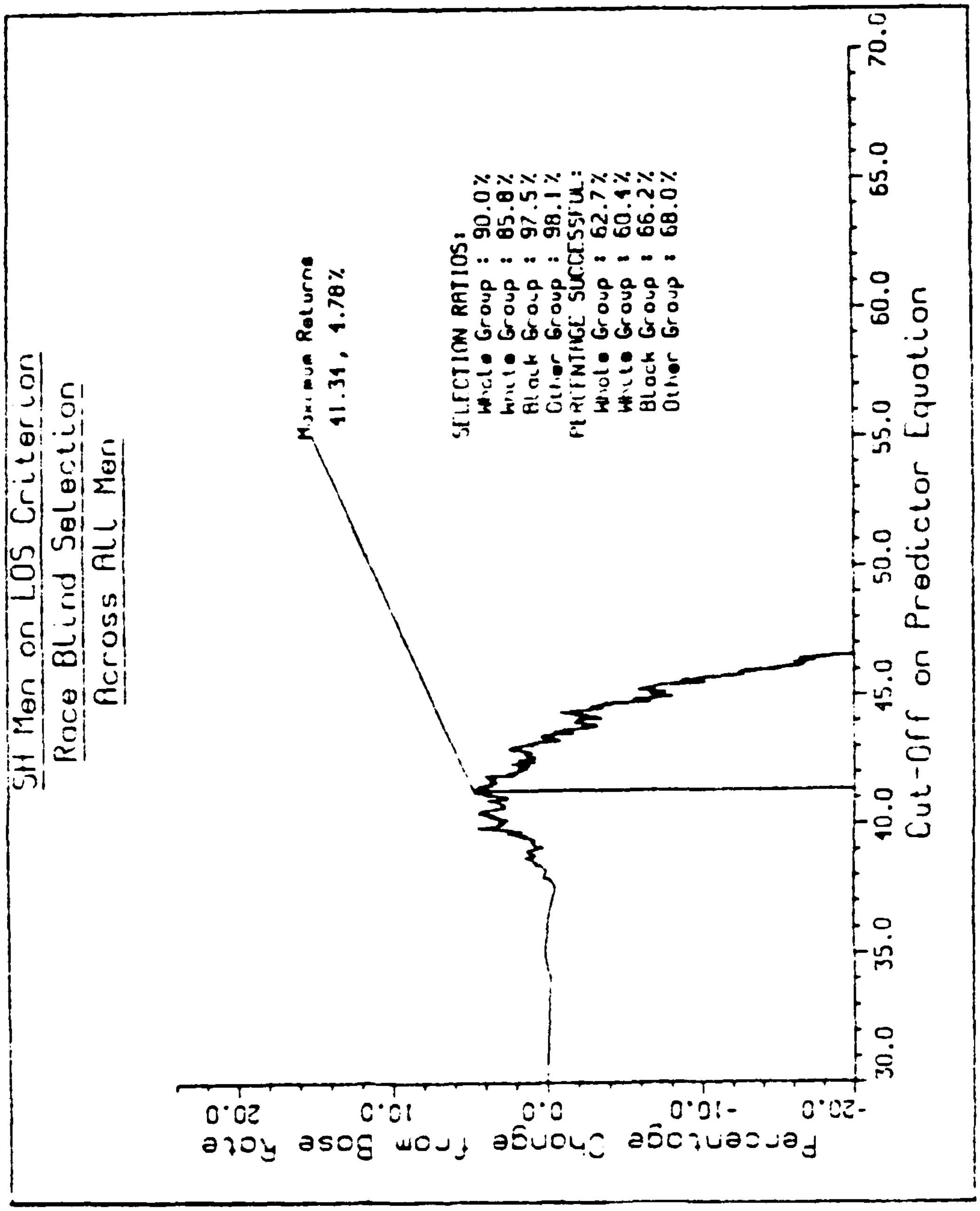


Figure 5.1 Race Blind Selection for SH on LOS.

The figures give several pieces of information, the most important of which is the line repesenting percentage change from the base rate, for various potential predictor cutting scores. For Figure 5.1, the base rate is determined as follows. Of the 2,020 male SH personnel 1,235 served four years or more, while 785 served for less than four years. Converting these values to probabilities gives a 61.14% chance for a male SH meeting the criterion of completing four years or more years of service, and 38.86% chance of failing it. Using the utility values given in Table IV and equation 4.1 we now calculate the base rate as the sum of the products of the chances of an outcome by its utility. The SH length of service base rate is 6,693.39. As the reader will recall, this is the value for selection set so low that all personnel are accepted. The first potential cutting score is the smallest predictor score from the sample of 2020. This score is 34.1208 and occurs for an SH who passed the four year criterion. With the cutting score set at this value, there are 1,234 sailors out of 2,020 who are successful and who would be correctly predicted as successful based on this predictive relationship. One sailor out of 2,020 would have been predicted to have failed the criterion for this cut-off, while 785 would have been incorrectly predicted to pass. The numbers are converted to probabilities to give 61.09%, .05% and 38.86%, respectively. These three probabilities have utility values of 19,260, -9360 and -13077. Summing across the products of probabilities and utilities gives a total value of 6679.23. Therefore, the return over base rate for a cut-off at 34.1208 is a -0.21% change. There is a slightly negative return over base rate for this cutting score. On Figure 5.1 the values 34.01208 and -0.21 are plotted as the first cut-off point and its respective return. The next highest predictor score, 34.4115, is choosen. This happens to be for a sailor who fails to meet the criterion. At a cutting score of 34.4115, the probabilities and their respective utilities are: 61.09% and 19,260; .05% and -9360; .05% and 13077; and, 38.81% and -13077. Summimg the product of these values gives a total utility of 6685.70, which is a -.11% decline over base rate. Therefore the next point plotted is 34.4155 and -.11. This process is repeated for all 2020 predictor scores keeping track of the value of the maximum return over base rate and its respective cutting score. Note in Figure 5.1 that values of base rate return which are less than -20.0% are not plotted.

Also shown in the figures are two other pieces of information. The selection ratio is defined as the percentage/proportion of personnel selected out of the total who apply. The values shown in the figures are selection ratios at the cutting score which maximise total return over base rate. On each figure selection ratio is shown either by race (for men) or rating (for women). The other statistic is the criterion proportions for those who are predicted to be successful. For Length of Service, the percentages who serve for at least four years are given, while for the other criteria the percentages by criterion categories are shown. This latter statistic gives an idea as to the expected performance of the group selected for this cutting score, while the selection ratio indicates how the selection device will impact the applicant group at this score.

1. ICS Criterion

As can be seen from Figure 5.1, the optimum cut-off score for the SH rating produces almost a 5% increase over the base rate utility. The predictor equation here is for male SHs based on all those variables selected for men in Table XII. The Figure shows the selection ratios and success percentages for the optimum cut-off score by race.

At this cut-off, white SHs have the smallest selection ratio of any race, but they also have the smallest percentage of success for those who would be selected. The majority of the 'black' and 'other' racial groups are selected and the highest percentages are successful in these groups.

Shown in Figure 5.2 are the cutting scores derived for race specific predictor equations. The best expected return is for the 'other' group (28.8%), followed by whites (8.9%), and blacks (1.7%). Set at these respective cut-off levels, 85% of the 'other' group are selected with about 76% of them serving beyond four years. For whites, 95% would be selected, with 59% successful. For blacks, virtually all (99.5%) would be selected, with a 67% success rate.

In comparing these figures for SH on LOS (Figures 5.1 and 5.2), several trends are apparent. The 'race-blind' selection seems to be most severe, in terms of selection ratio, on the white group. When the white specific predictor equation is employed, more whites are selected, for a slight decline (about 1.5%) in the percentage success. For the black group, a race specific predictor also improves the selection ratio, with a slight decline in the success rate. However for the 'other' race group, the race specific predictor equation markedly reduces the selection rate (down by about 13%), but boosts the success percentage by almost the same amount.

Figures 5.3 and 5.4 show LOS results for the PN rating. Race blind selection produces about a 6% improvement over base rate. As for SHs, the race blind predictor is most severe on the white group in terms of selection ratio. However, for those whites selected, their success percentage is better than the black PNs. While the blacks are the least successful for this predictor equation, about 97% of them are selected. At the cut-off level for this race blind equation, 100% of the 'other' racial group is selected, with about 76% of them being successful.

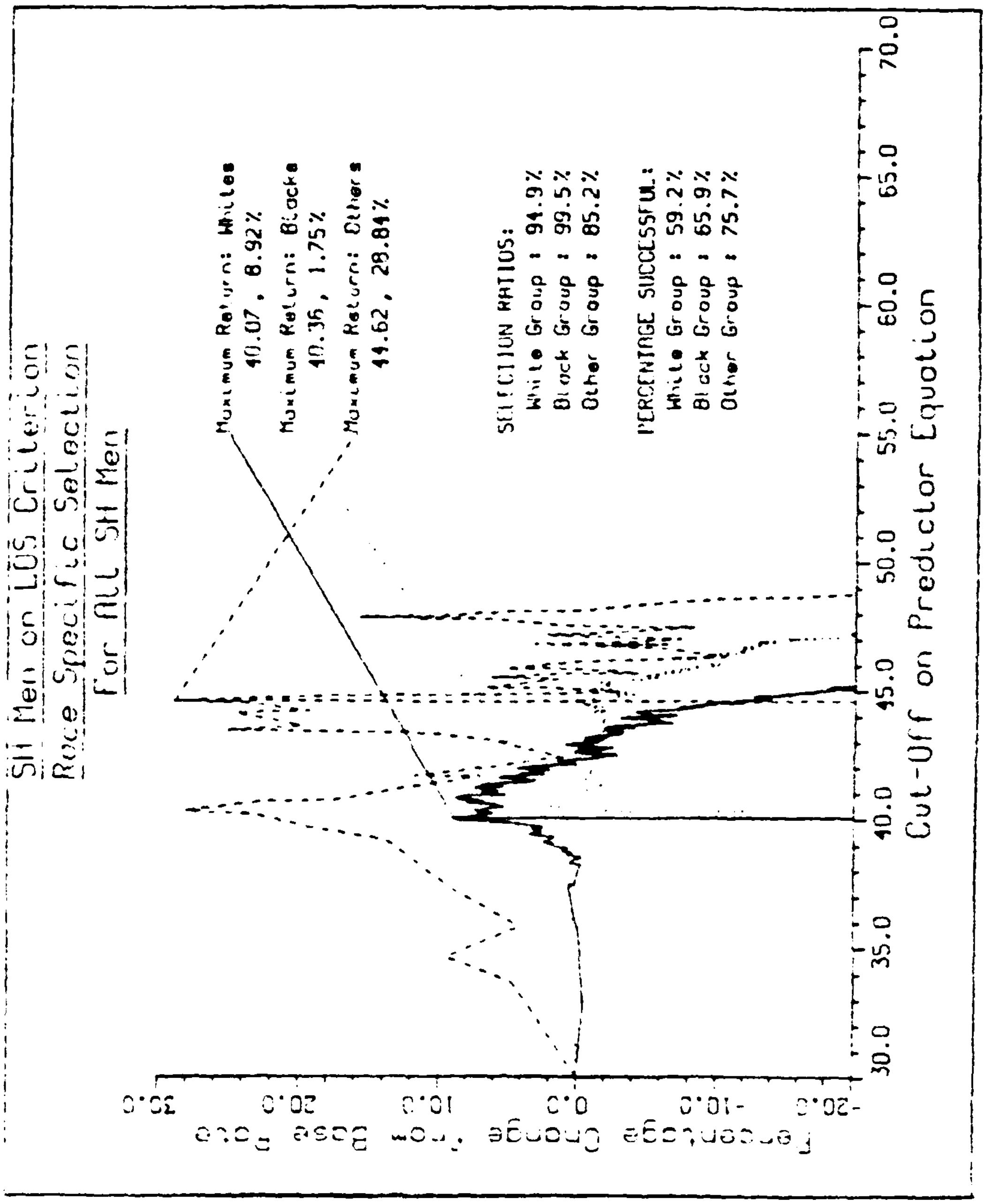


Figure 5.2 Race Specific Selection for SH on LOS.

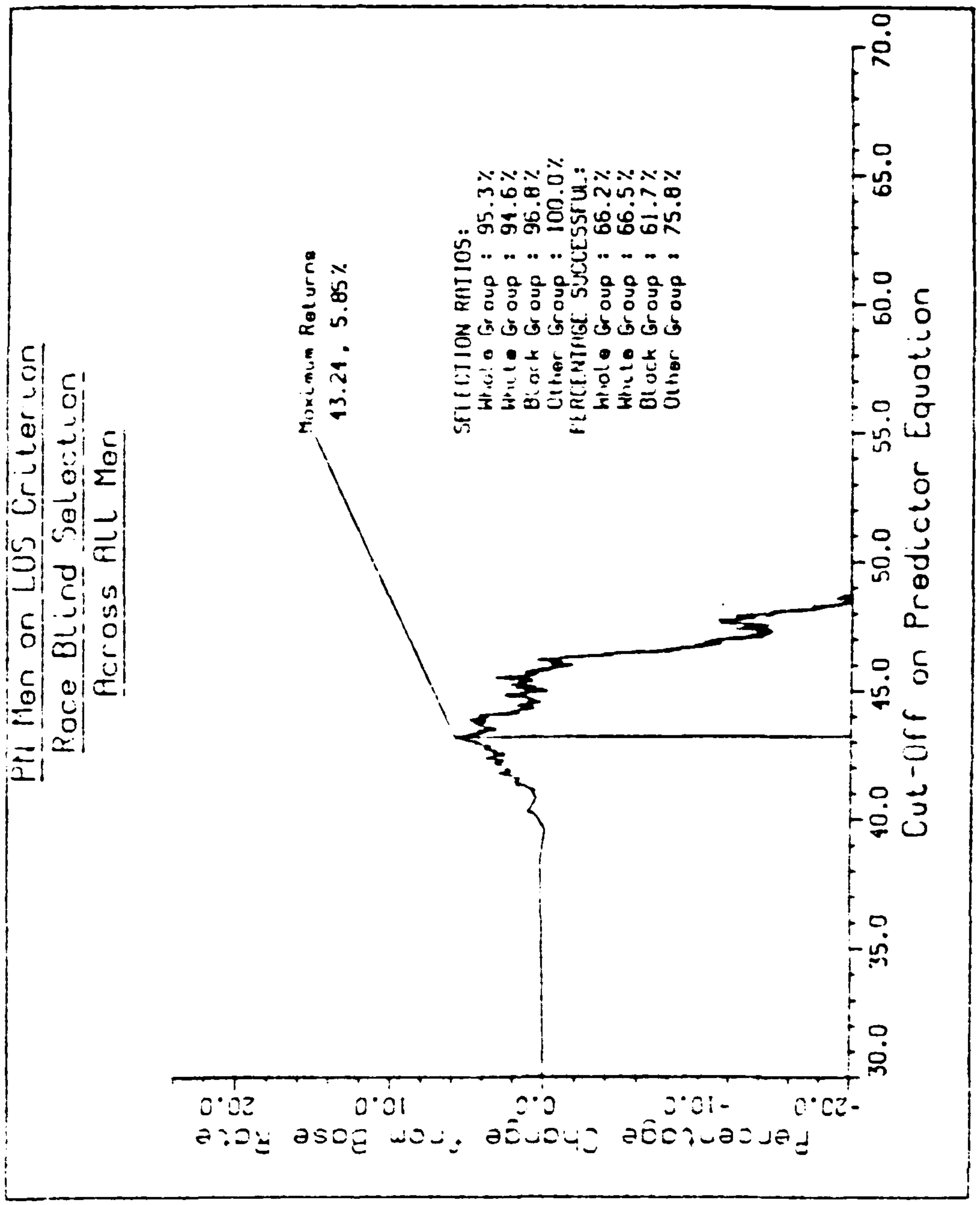


Figure 5.3 Race Blind Selection for PN on LOS.

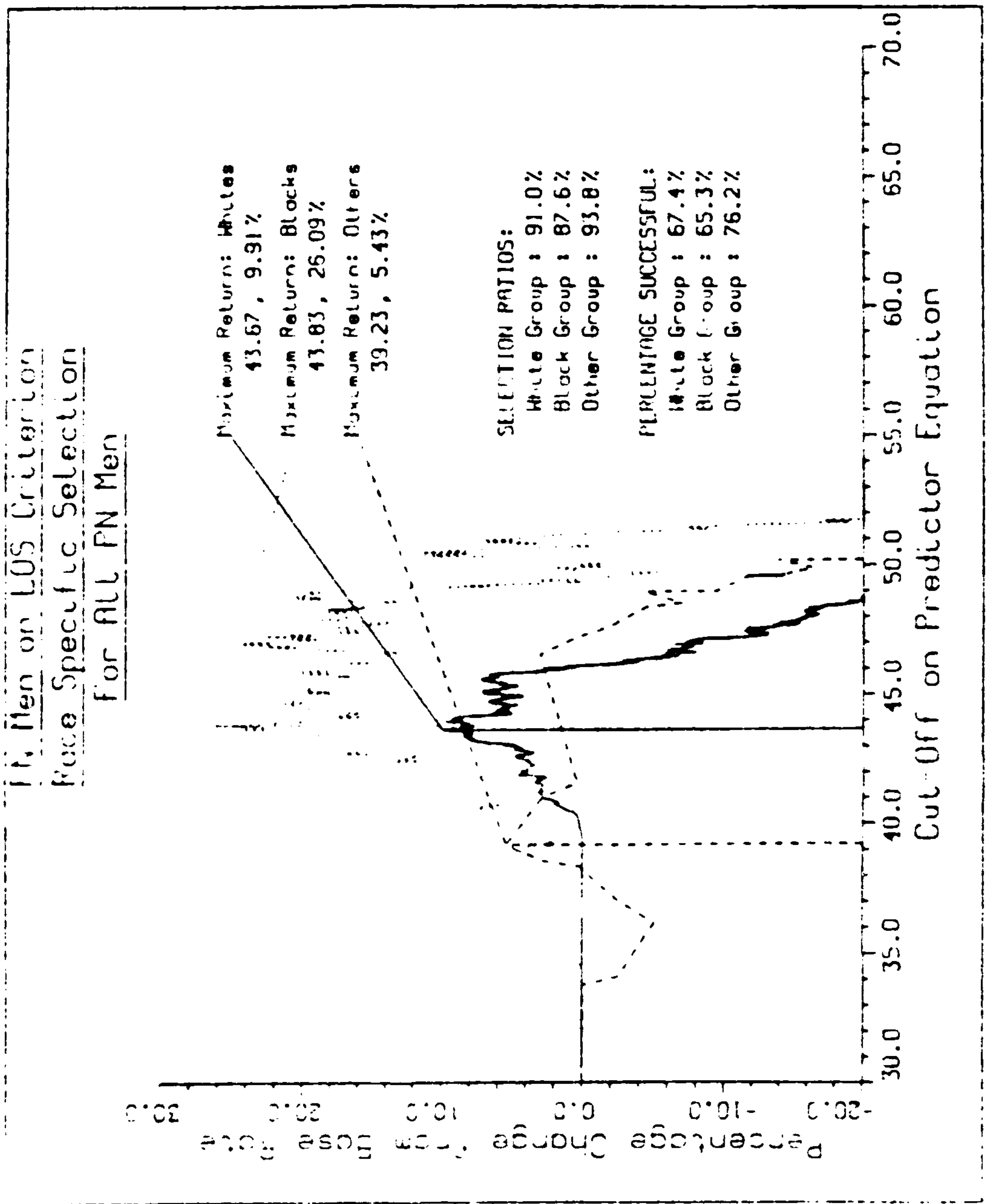


Figure 5.4   Race Specific Selection for PN on LOS.

The return for individual race selection equations (Figure 5.4) is greater than from the race blind selection for each race. Blacks show the biggest improvement, with a reduction in their selection rate, down from 97%, but an improvement in their success percentage. Next come whites whose selection ratio has also been reduced but their success percentage improved. This is followed by the 'other' racial group, who are now selected at a rate of about 94%, with a slight increase in overall success percentage. Therefore, for all racial groups, race specific predictors have lead to a decline in selection rate, but improvements in overall utility and success percentages.

As can be seen in Figures 5.5 and 5.6, the same trends are not true for ATs on LOS. Race blind selection, lead to about a 3% improvement on base rate, with all races being fairly close together on selection ratios and success percentages. The white race is marginally ahead on both of these indices. In Figure 5.6, about the same values apply for whites on race specific selection. However, the race specific equations for blacks and for the 'other' racial group have determined optimum cutting scores for which all of the cases would be rejected. That is, no one is selected and thus no one is successful. This result may have been expected for AT 'others' because of the insignificant validity coefficient for this group (see Table XXI). Low validity suggests that any optimum cut-off value is likely to be spurious rather than meaningful. However, Table XXI shows a highly significant validity for blacks. Therefore deriving an optimum cut-off which rejects everyone is certainly not expected.

Shown in Figure 5.7 are the LOS results for white women PNs and ATs. Both optimum cut-offs have improved the base rate figure (24% for PN and 36% for AT). While the selection ratios are lower than their male counterparts, so

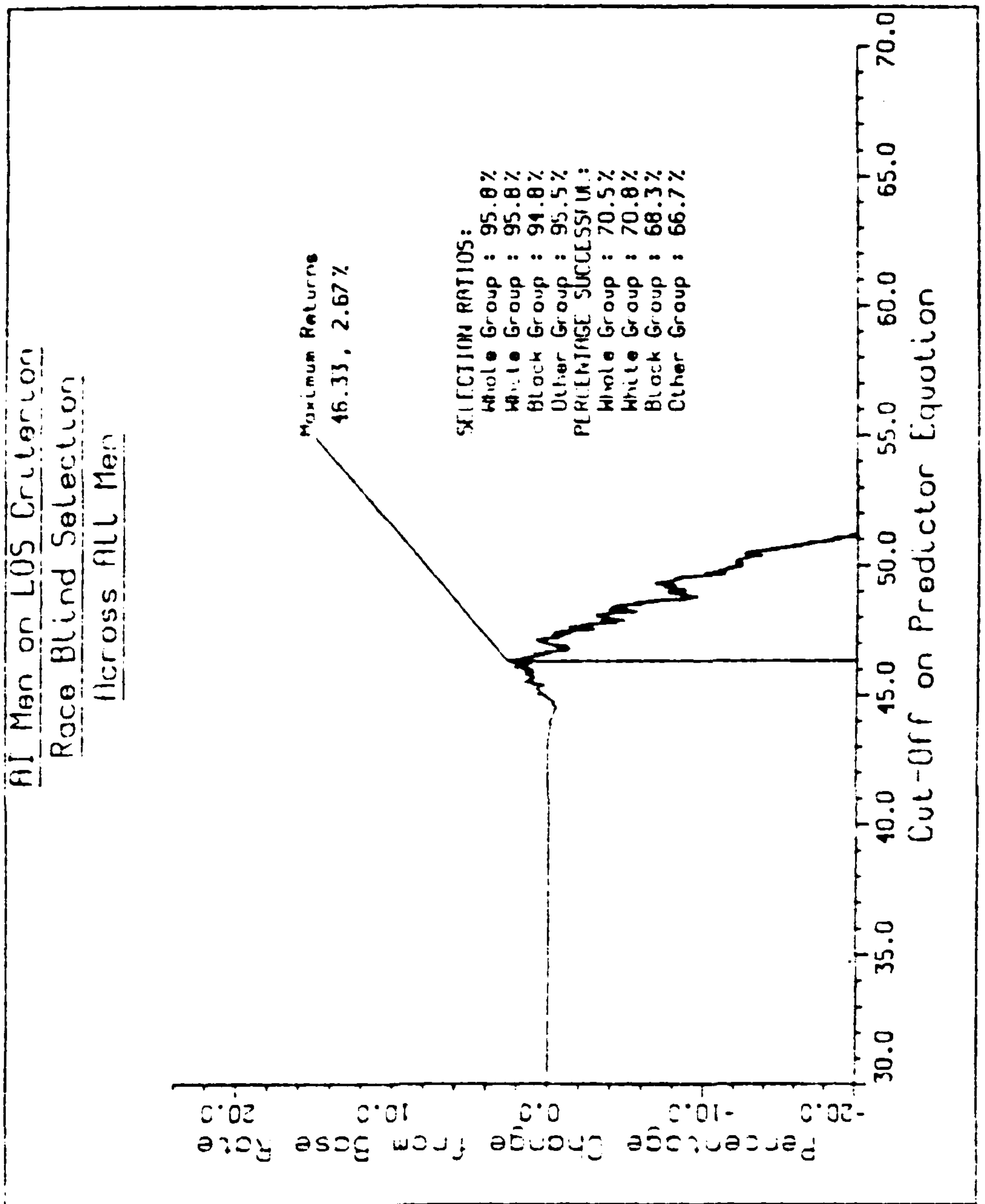


Figure 5.5 Race Blind Selection for AT on LOS.

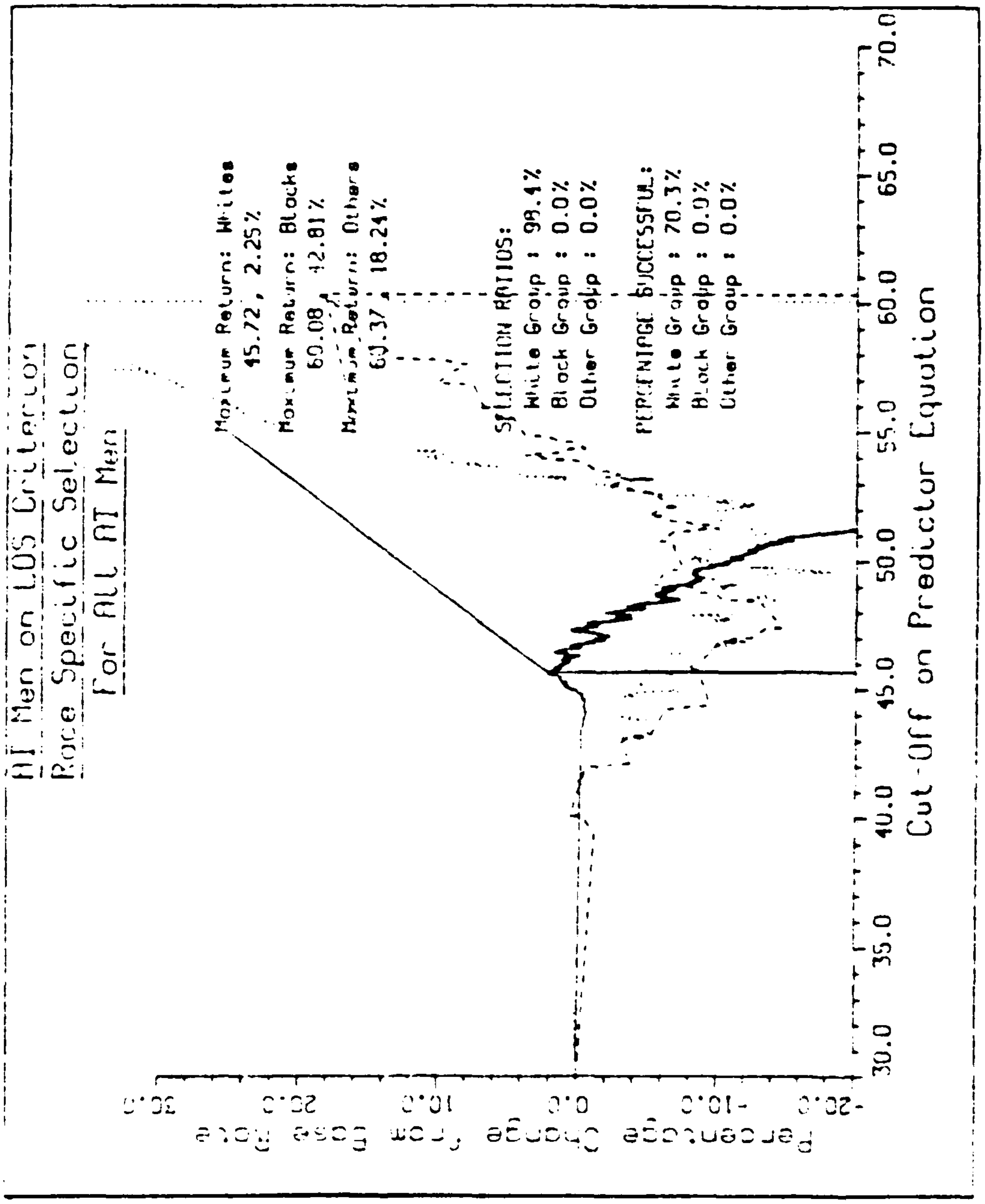


Figure 5.6 Race Specific Selection for AT on LOS.

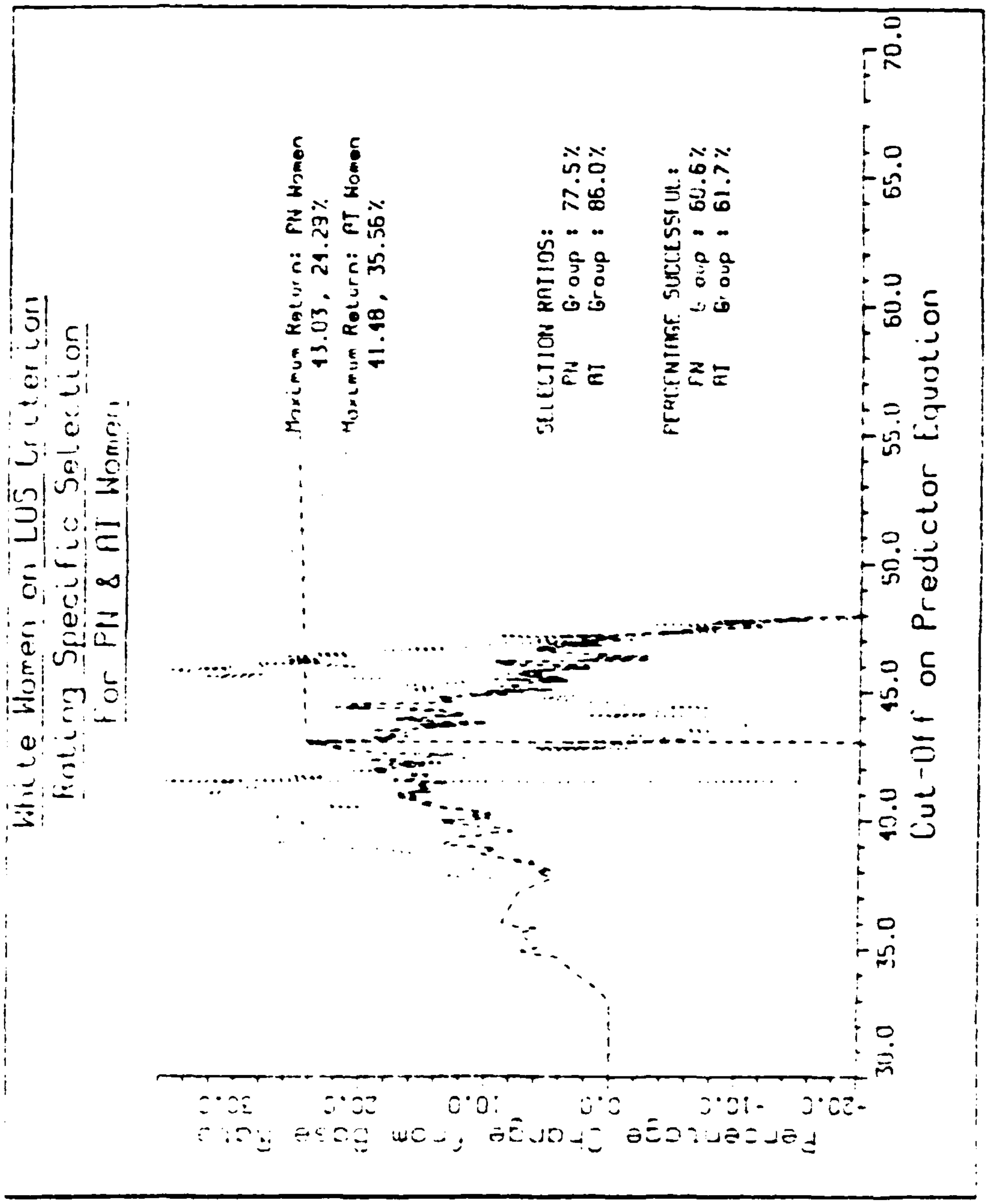


Figure 5.7 Rating Selection for PN & AT Women on LOS.

too are their success percentages. This is not so surprising, since Tables X and XI show women have the smallest average LOS of all groups in these ratings.

2. Goodguy Criterion

Shown in Figures 5.8 to 5.14, are the results of the analyses on this criterion.

For the SH rating, race blind selection leads to a selection ratio for whites that is the smallest of all three races. However, as can be seen from Figure 5.8, the optimum cut-off yeilds a return of just over 2%. For race specific selection (see Figure 5.9), the returns for whites and 'other' races are well over 30%, but in this case, only 65% of the white group would be selected. The selection ratios are reduced for all races, compared with race blind selection.

The returns for the PN rating on the goodguy scale, follow a similar pattern as can be seen from Figures 5.10 and 5.11. However, for this rating, blacks in the race blind selection, have the smallest selection ratio, but they also show the largest return (42%). Indeed, race specific selection has improved their position in selection ratio terms from the lowest for race blind selection to the mid-point on race specific selection.

For ATs on the goodguy scale, race blind selection does not produce any positive return over base rate (see Figure 5.12). The return for race specific selection (Figure 5.13), is substantial for blacks and 'others' (38% and 76%, respectively) but a little more than 1% for whites. For the 'other' race group, this high return is achieved with a selection rate of about 61%. For female PN and AT on the goodguy criterion, see Figure 5.14, the returns are almost 600%. The cut-off for AT seems unrealistic, since only 14% are selected, whereas for PN the selection rate is 75% means that a reasonable proportion will be selected.

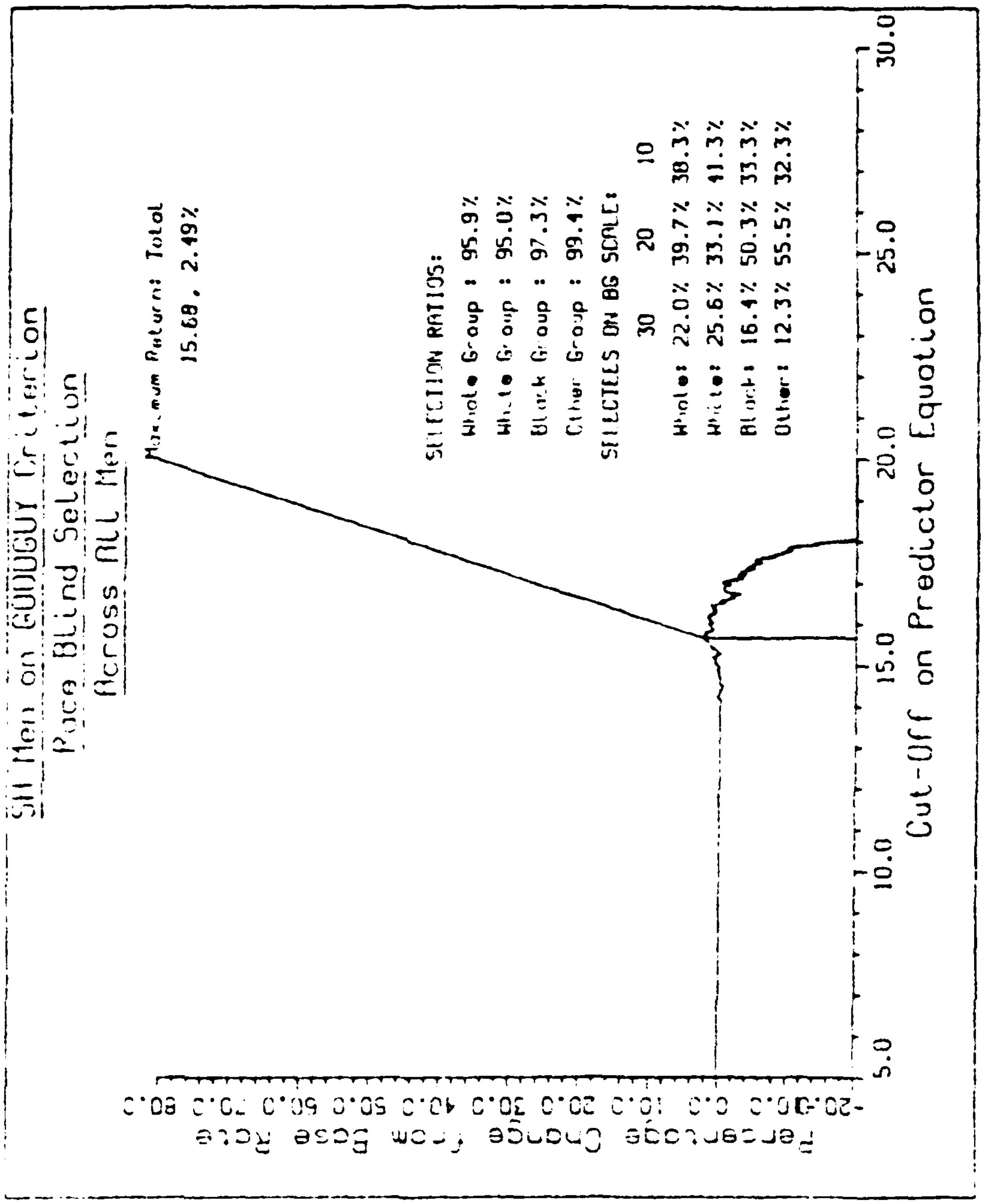


Figure 5.8 Race Blind Selection for SH on Goodguy.

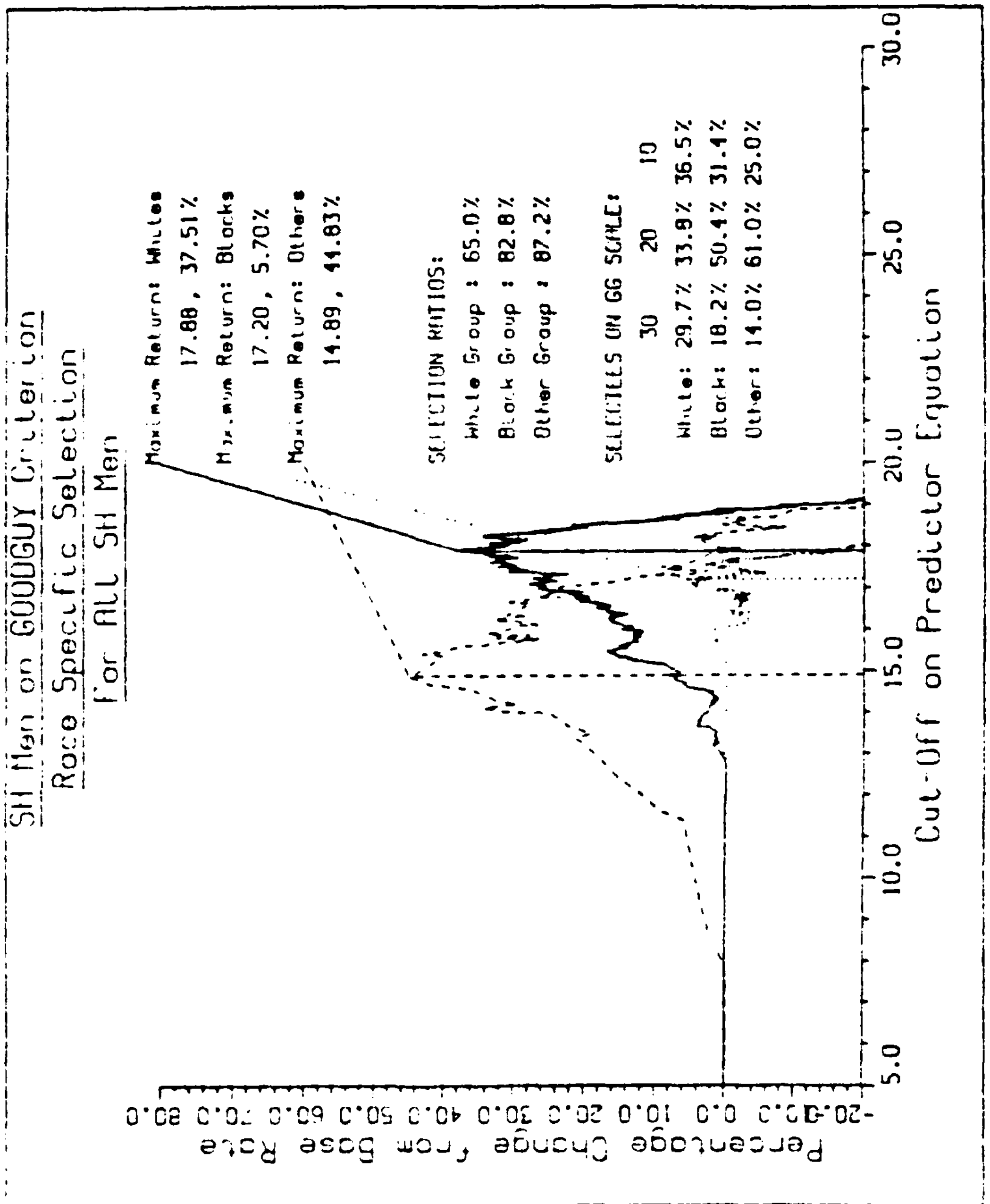


Figure 5.9 **Race Specific Selection for SH on Goodguy.**

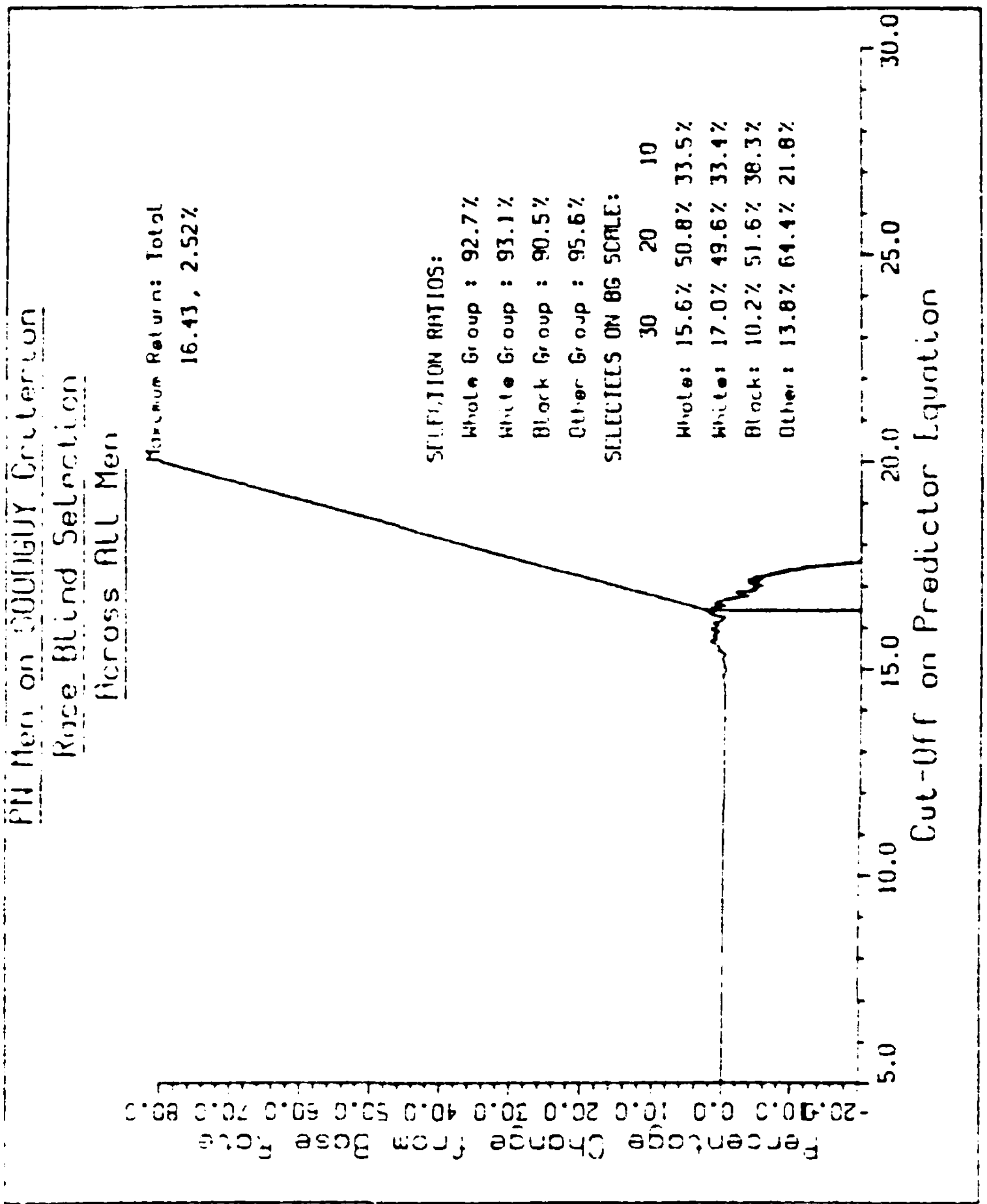


Figure 5.10 Race Blind Selection for PN on Goodguy.

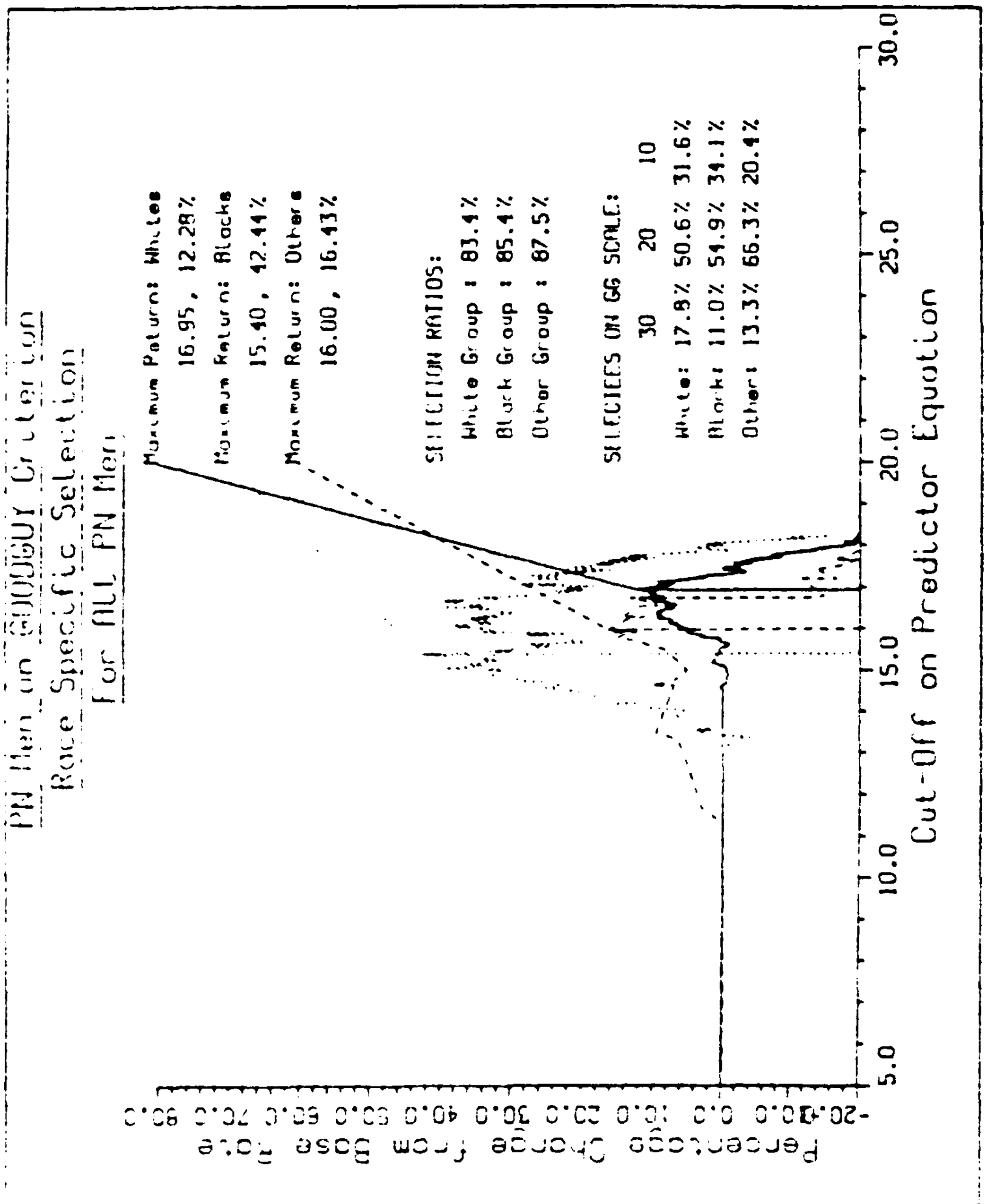


Figure 5.11 Race Specific Selection for PN on Goodguy.

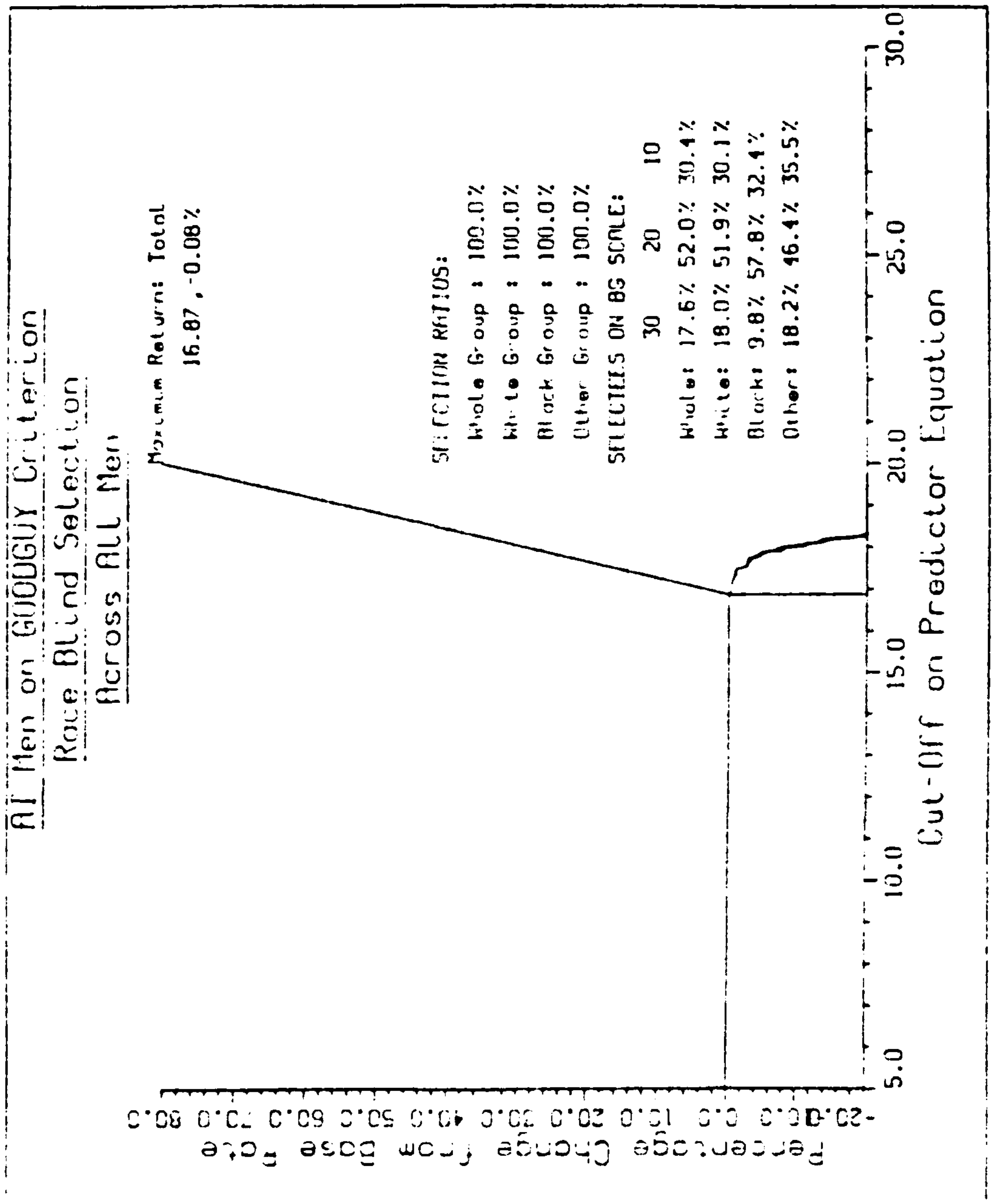


Figure 5.12 Race Blind Selection for AT on Goodguy.

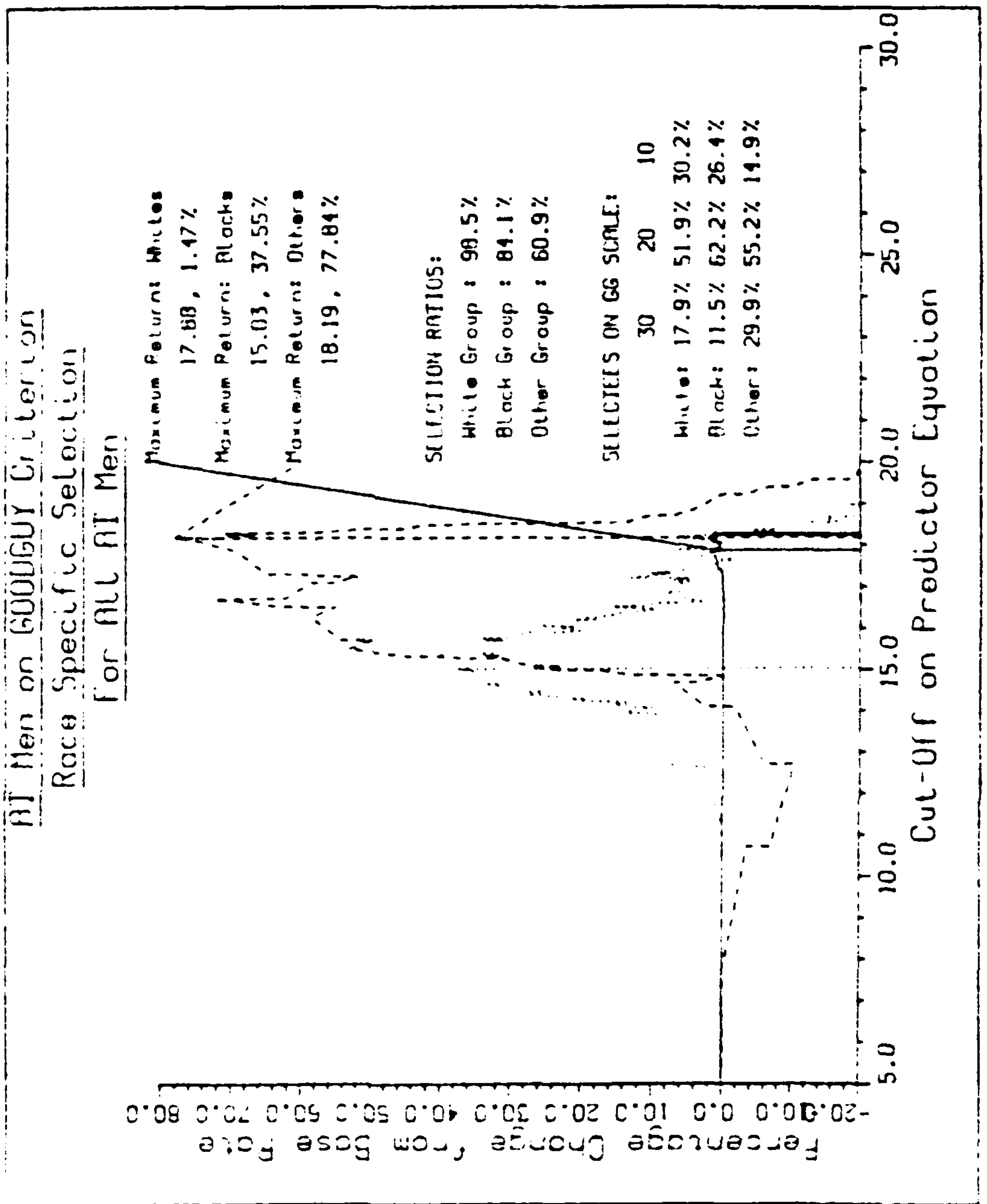


Figure 5.13 Race Specific Selection for AT on Goodguy.

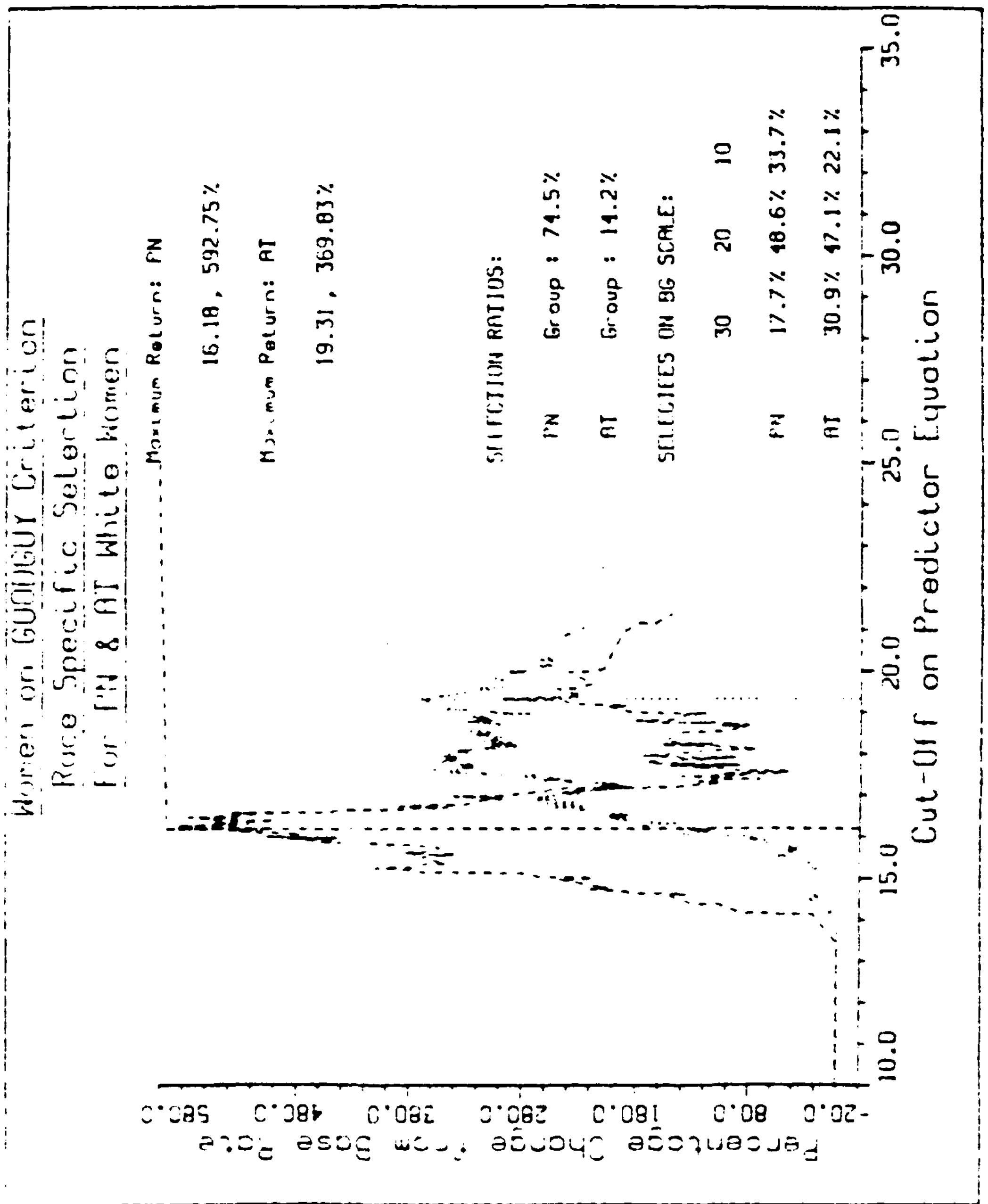


Figure 5.14 Rating Selection for PN & AT Women on Goodguy.

3. Badguy Criterion

The seven Figures 5.15 to 5.21, show the results for this criterion.

Race blind selection for male SHs, see Figure 5.15, shows about a 3% improvement over base rate. White men are the race which have the lowest selection ratio in race blind selection. Moderate improvement for whites and the 'other' racial group occur for race specific prediction (Figure 5.16).

For PN, see Figures 5.17 and 5.18, there is virtually no effective return using the badguy criterion. While the 'other' racial group shows only about a 6% improvement over base rate, the selection ratio of 98.2% is extremely high.

As can be seen in Figures 5.19 and 5.20, for AT men no selection device has been found that produces an improvement over the base rate.

For women on this criterion, Figure 5.21 shows a positive (AT) and a negative (PN) return. The AT return is difficult to interprete and results in a 99% selection ratio.

1.0 1.1 1.25 1.4 1.6 1.8 2.0 2.2 2.5 2.8 3.2 3.6 4.0 4.5 5.0 5.6

MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS - 1963 - A

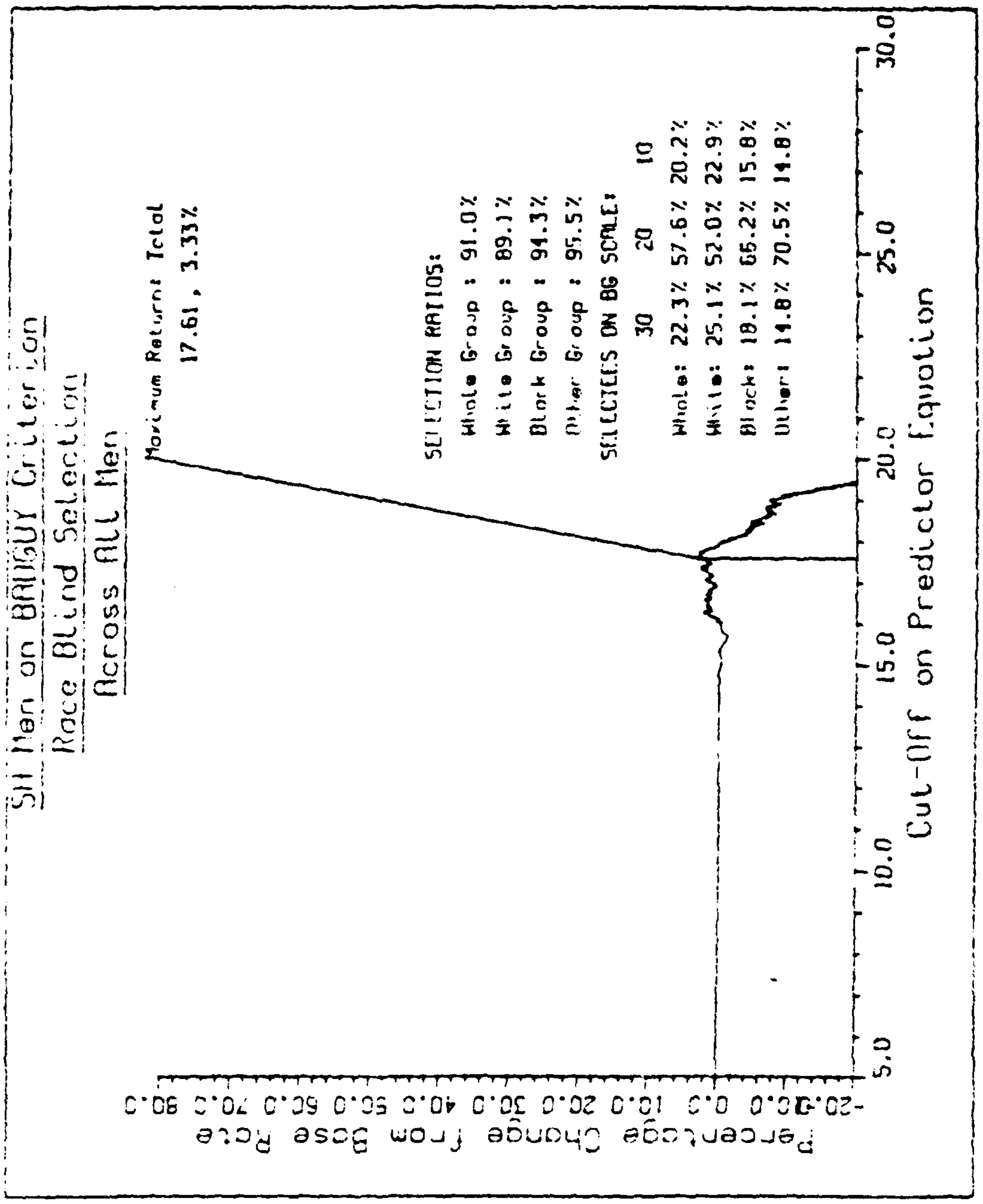


Figure 5.15 Race Blind Selection for SH on Badguy.

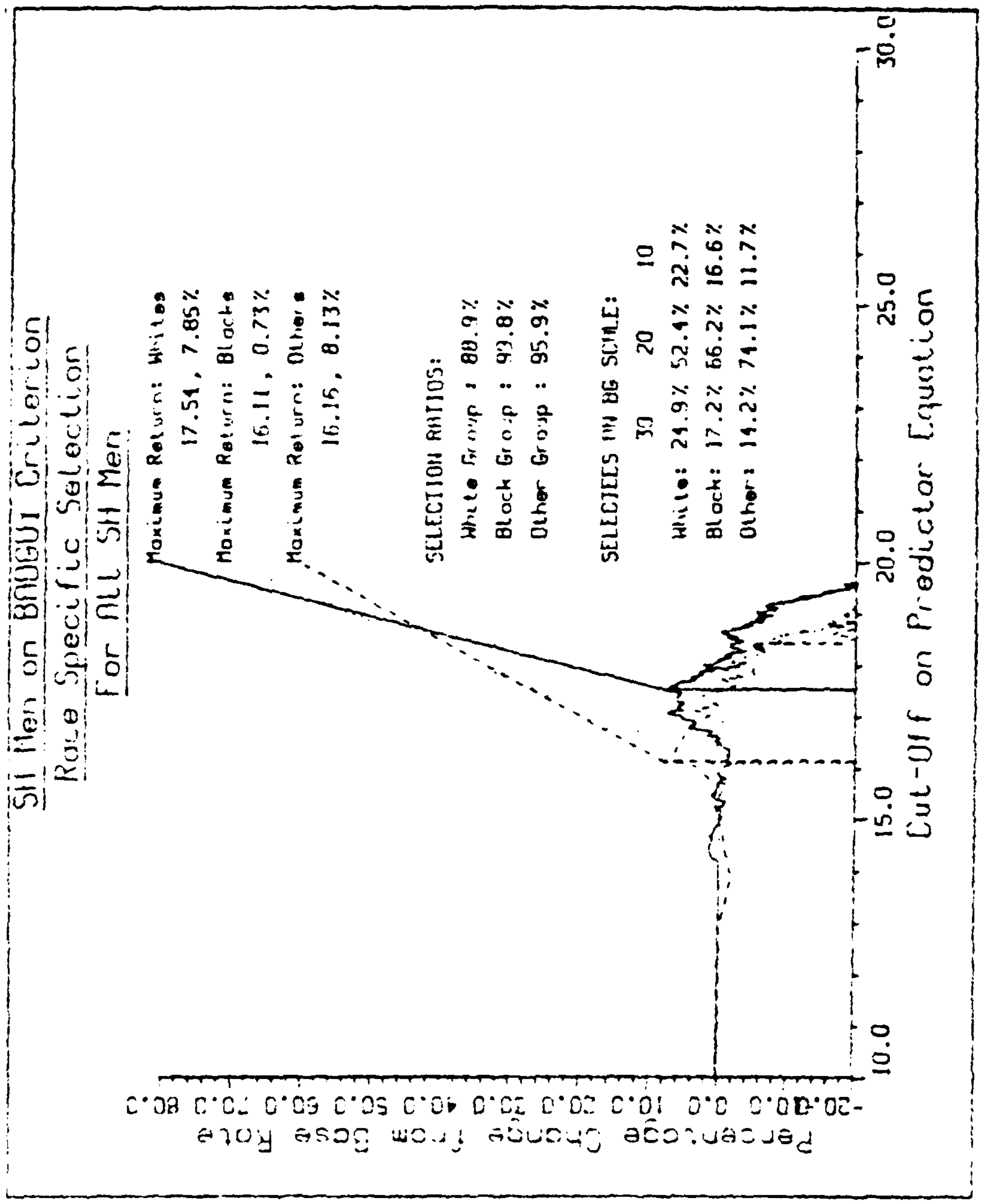


Figure 5.16 Race Specific Selection for SH on Badguy.

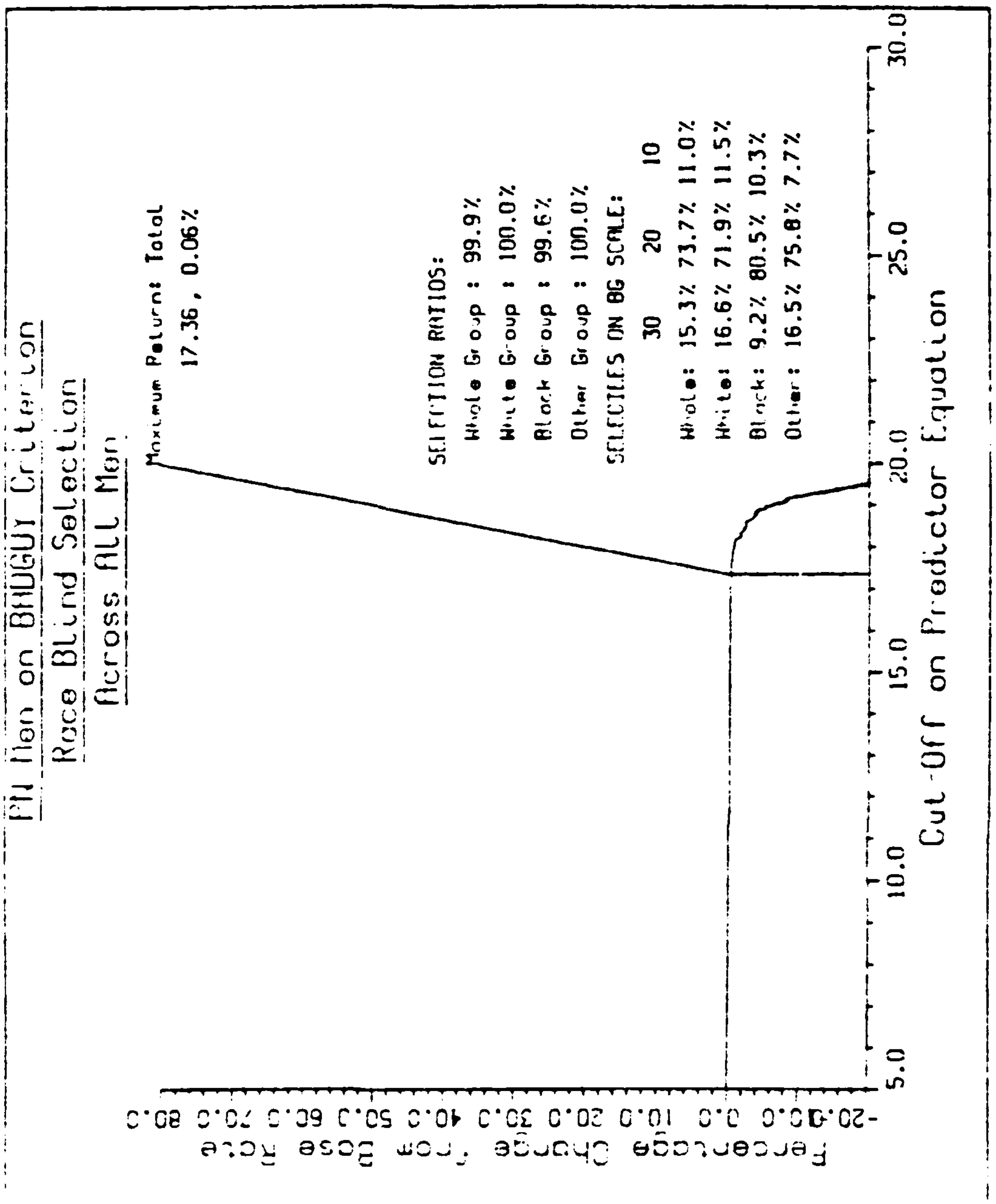


Figure 5.17 Race Blind Selection for PN on Badguy.

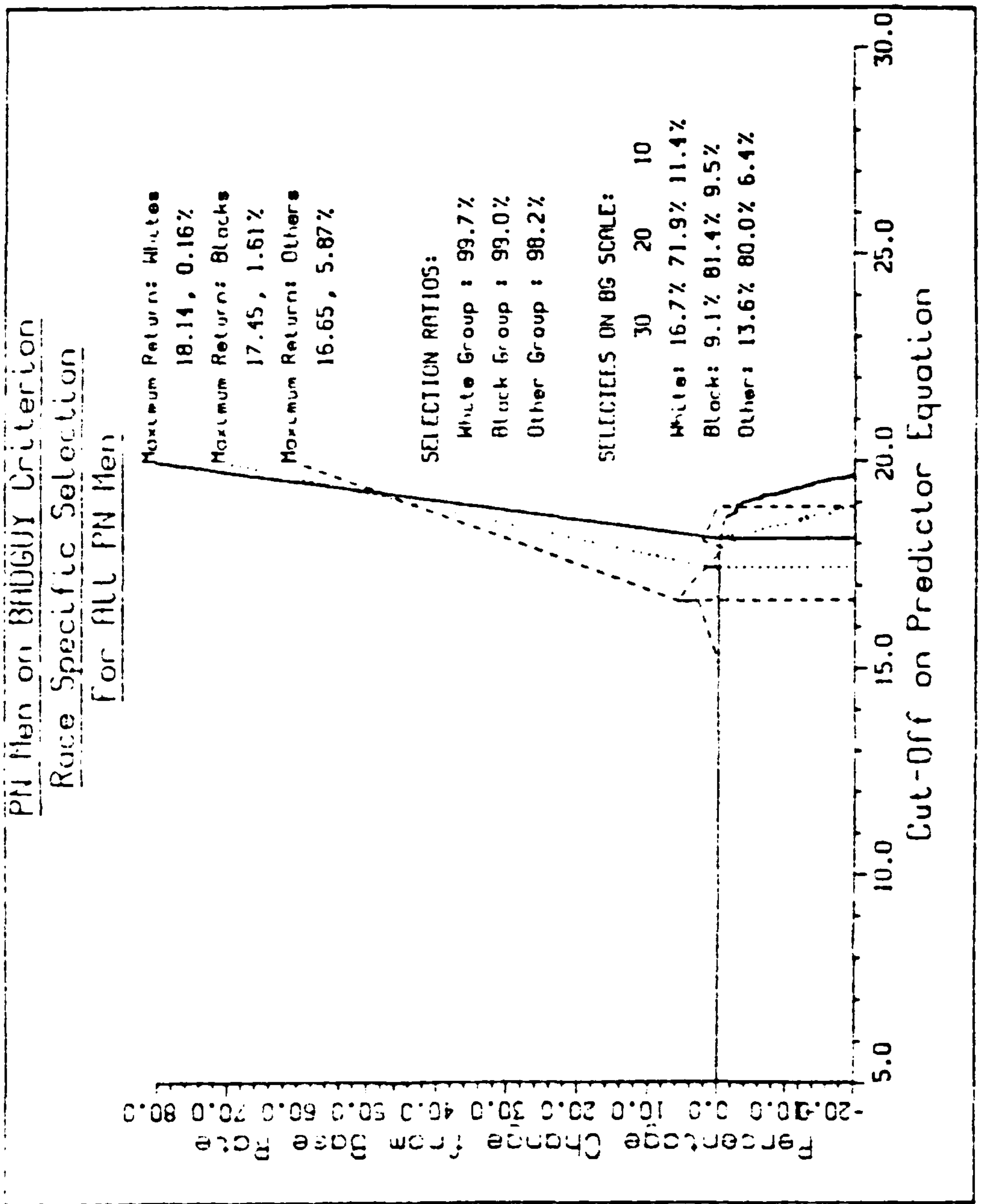


Figure 5.18 Race Specific Selection for PN on Badguy.

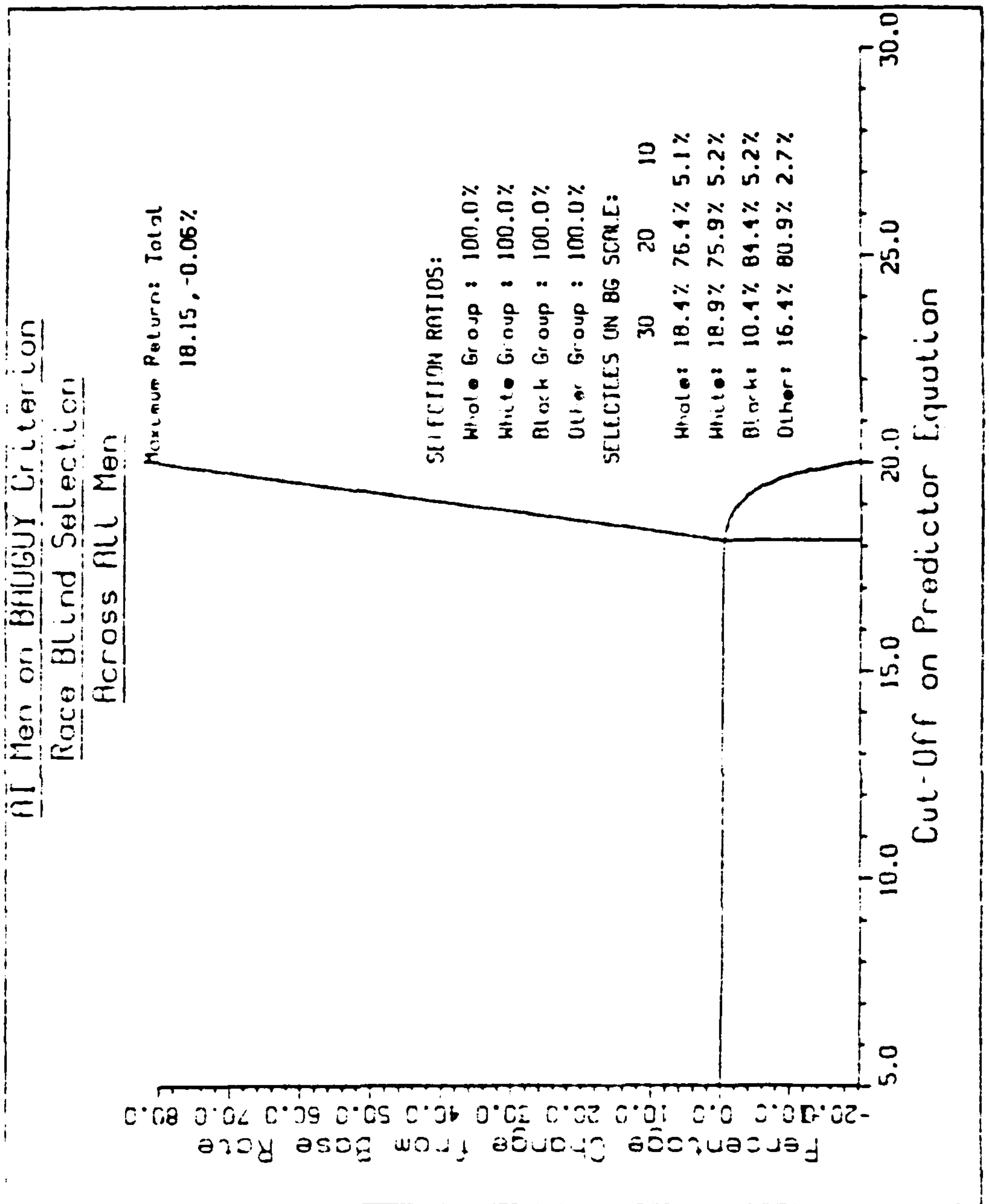


Figure 5.19 Race Blind Selection for AT on Badguy.

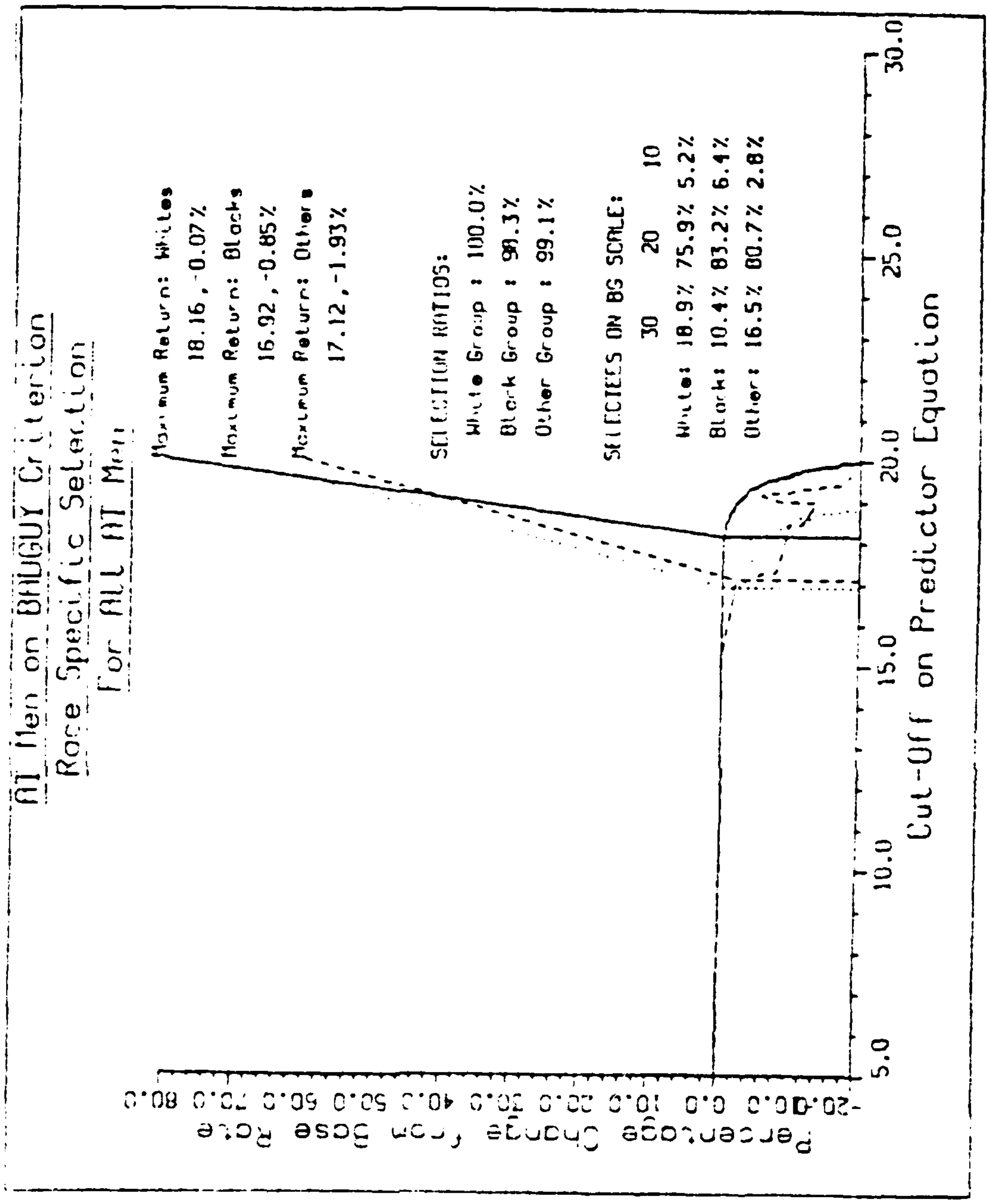


Figure 5.20 Race Specific Selection for AT on Badguy.

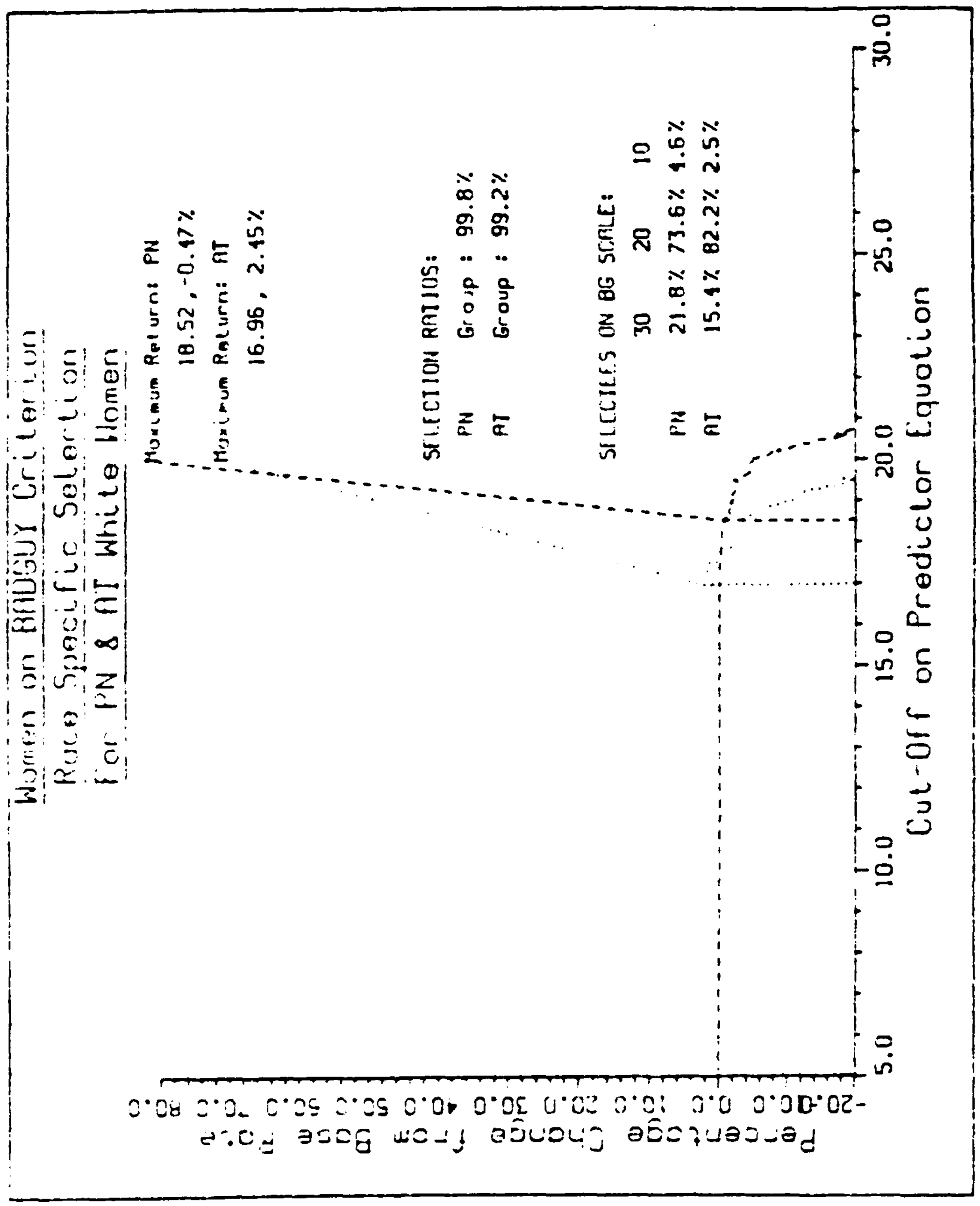


Figure 5.21 Rating Selection for PN & AT Women on Badguy.

VI. DISCUSSION

This chapter is divided into two sections. The first deals with the results in relation to the research hypotheses. The second discusses other findings.

A. RESEARCH HYPOTHESES

The stepwise regressions, particularly on the LOS criterion, tend to support the first hypothesis which predicted that entry age, education level and ability tests would be significant predictors of performance. Entry age was not selected for every regression, but it proved to be significant for at least one of the race/sex groupings on LOS for each rating. The relationship between age and the criterion was always positive: longer service is indicated for greater age at entry. Educational level was often selected too. However the direction of the relationship was not as consistent. For some groups it was positive and for others negative. There seems to be a trend for the coefficient for education level to be positive in low complexity ratings and negative in the higher complexity ratings. This finding could be a related to restriction of range problems which were noted in the literature review. At least some of the ASVAB scores entered every stepwise regression. The signs were not consistently positive or negative.

At least for one criterion (i.e. LOS) the validity coefficients when averaged across groups, show a trend to increase with job complexity, a relationship predicted by the second hypothesis. On the whole, the validity coefficients are large enough to suggest that the predictor equations are sufficiently powerful to improve selection on these three criteria. Comparing criteria, LOS has the highest validity.

B. OTHER FINDINGS

The descriptive results yeilded several general findings. They seem to support the view that whites have superior performances on psychometric tests of ability. On each on the ASVAB subtests, with one or two exceptions, the white group in each rating has performed better, in terms of raw score, means than the other two race groups. On the other hand, this race, for each rating, is younger at entry and (perhaps therefore) has the lowest educational level, with the smallest proportion being married variables which might be expected to be associated with lower ASVAB scores. For the SH and PN ratings, the white group has a shorter LOS, but with the highest proportions of in-service 'good' and 'bad' behaviours. For the AT rating, blacks and whites have about the same LOS (blacks serve slightly longer), and whites are more likely to have positive goodguy performances. The white women have similar performances as white men on both the predictor and criterion variables, except that they have the shortest LOS and generally exhibit the smallest percentage of 'bad' behaviours.

The estimation of cut-offs, for all criteria, has supported the notion that the significant relationship between predictors and criteria can be useful in a selection setting. As measured against the concept of base rate, many groupings within ratings can be better selected.

The findings of this reseach support the current US Navy selection procedures. For every stepwise regression in which it was selected, the SCREEN score (which is derived directly from the current procedures) was always positively related to the criterion.

From the averaged validities it appeared that the LOS predictors provide a means of increasing the selection utility across most trade, race and sex groupings. However, when optimum cut-offs were derived it was apparent that virtually none of the AT relationships would yield any reasonable utility improvement over the base rate. For the other ratings it appears that, except for SH blacks, there are positive returns to be obtained through selection based on the identified predictor/criterion relationships. The highest returns are for PN and AT women. This is despite the fact that some predictor equations had validities which were small, and in some cases statistically insignificant (e.g. PN 'other' and AT women). Apparently, although the validity coefficient was insignificant the relationship was still strong enough to yeild a cutting score which improved overall utility.

The other two criteria, although they did not have double cross-validition coefficients which were as large as for LOS across groupings of personnel, also apparently are predictable from pre-entry information. However, when the cut-offs were determined, the goodguy criterion appeared to have potential for improving selection in all groupings, withthe exception of AT white women. For only three groupings on the badguy criterion was a cut-off found which returned better than 5% over base rate. In a real sense these seem more useful criteria than LOS, since they appear to capture more relevant information relating product and/or cost. The extent that personnel are promoted during a period of service and the extent to which costly negative behaviours can be avoid are significant contributers to the organisations overall effectiveness. It would appear that the decision centered validity approach has revealed a number of potentially useful relationships which may be used for better selection.

The race blind selection cut-offs provided some interesting findings. For the SH rating in particular it seems clear that the white race would have greater proportions rejected than the other two racial groups if the race-blind prediction device was used. However, as has already been mentioned, more of the white SH personnel should be rejected since they exhibit poorer criterion performances than the 'other' races. For the PN rating, race blind selection also selects the smallest percentage of whites. This is a more curious result since for both ratings, most of the personnel were white and one would expect that they would dominate the regressions to the extent that greater proportions of whites would be selected than would the other races. For the specific selection the white selection ratio at the optimum cut-off is no longer the lowest for the LOS, but it is the goodguy scale. It appears that whites do not perform as well as the other races in these ratings (PN and SH).

The decision centered validity approach has turned out to be a powerful tool for the present research. It has allowed a means of directly demonstrating the usefulness of a composite predictor in a future selection role. Provided appropriate utility values can be derived, it is obvious that the methodology employed in this study could be broadened. For example, in evaluating which cut-off to apply, the criterion can be segmented into as many separate categories as is required to reflect the different individual returns for the pairs of predictor outcomes. At one extreme in criterion categorisation is the LOS methodology employed here, in which the criterion was considered as a dichotomy. At the other is the situation in which each and every subject in the research could be individually evaluated in utility terms. For example, for every subject a value could be placed on the service he/she provided, and an estimation could be made as to the value to the Navy (positive or negative) in the case that this person was not enlisted because of selection screening. These two values, which would probably not be the same, could be taken as predictor accept and predictor reject utilities and then used in exactly the same manner as in this study to determine an optimum predictor cutting-score which maximises total utility for all subjects. If the utility values were accurately derived from costs and products, then it may be possible to express selection utility in dollars per enlistee, which would be a readily acceptable means of justifying a selection procedure.

In applying the regression equations developed here to an actual setting, there are two other important considerations. These are both related to decisions that the Navy would make with respect to implementation. In this thesis, there was no external Navy decision as to which criteria of performance is the most useful for personnel selection. Length of service is widely used as a measure of good/desirable behaviour, but for some employments, it may not be so important. Indeed for some employments, "short" tenure may even be a goal (e.g. those employments which serve primarily to prepare personnel for later jobs). So before any attempt is made to implement these results the criterion would need to be specifically defined. It could be that there are other more pertinent criteria than those employed in this thesis.

The other issue about implementation is related to costs and has two aspects. First, the present study estimated costs and utilities through published cost data. These figures were clearly and directly relevant for the LOS criterion, but even in this application assumptions were made as to the relative value of the predictor/criterion outcomes. There are, perhaps, better accounting or economic procedures to estimate these values than was used here. An evaluation of the utilities employed here is needed before implementation of these selection devices. Second, the other aspect of cost is directly related to implementation. In determining the optimum cutting-score on the predictive device for this study and expressing this as a percentage return over base rate, it is clear that it does not take into account any costs associated with implementation. A cost-benefit type of analysis would be necessary to weight up the total costs of implementing and maintaining the procedures, against the expected return in terms of more profitable selection. The data presented here are relevant to this type of analysis, but returns over base rate, for example, would need to directly take account of the numbers upon which each of the graphs in the previous chapter was based on. Some of the very high returns over base rate may be seen in a different light when the numbers of cases is taken into account. Base rate is a utility per enlistee: if the applicant pool is small, total utility (the product of the number of enlistees by utility per enlistee) might not be so high. If the cost of implementing and maintaining a procedure was very high, then the expected return over base rate would have to be even higher, so as to produce a positive return.

VII. SUMMARY AND CONCLUSION

This thesis set out to develop selection standards for three US Navy ratings which varied in terms of their complexity requirements. The selection literature was extensively reviewed. This gave direction as to the variables likely to be predictive of subsequent performance as well as some guidlines for choosing appropriate criteria to predict. The method of evaluating the predictors in forecasting subsequent performances took two forms. The first was the traditional method for estimating validity coefficients through double cross-validation. The second applied utility estimates to various predictor/criterion outcomes so as to determine the cut-off score on the composite predictor which maximised overall utility. These analyses were done controlling for the effects of such moderator variables as job complexity, race and sex.

Within each rating a number of useful predictor/criterion relationships have been found. These relationships were shown to be valid and cutting scores derived which maximised total utility of the selection device. Some results suggest a significant race component in performance. This finding suggests the white race does not perform as well as the other races in the SH and PN rating. The signs on some ASVAB coefficients indicate negative relationships with criterion scores. This result was not expected.

This research has illustrated the relationships that exists in the data and the use to which they could be put for selection purposes. Implementation of any of these relationships as a selection tool requires a confirmation that the criteria used here are relevant and useful to personnel managers in the US Navy. In addition, the current analysis has been performed without any consideration of the possible costs of the introduction and maintenance of the selection procedure. It would be necessary to conduct a cost benefit analysis to determine if the total expected returns from implementation were sufficient to compensate for the costs involved and therefore to produce a positive total return. To do this properly requires Navy decisions about the appropriateness of the criteria to be predicted and the utility and disutility values to be assigned to correct and incorrect predictor/criteria classifications. Such a cost benefit analysis has not been attempted here, but is relevant for future research.

APPENDIX A

EXAMPLE PROGRAM LISTINGS BASED ON PROCESSING SH

This appen ix gives sample listings for the programs used to analysise the SH rating. These exactly parallel analyses for the other ratings.

1. SAS Program to Select SH from NPS Cohort Data

On the following pages (Table XXVI) is a listing of the SAS job to select out the SH from the Cohort File of 206,000 and to place their records on Mass Storage.

The variables are read in from raw data and the first two pages of Table XXVI, show the variables locations and their descriptions. This lists all the variables available in the cohort even though only a smaal number of them were actually used for the present analysis.

The reader is referred to the SAS "User's Guide" for an explanation of the procedures involved. Comments throughout the listing explain some of the program steps.

2. SAS Program to Perform Regressions

The following two tables give listings for two SAS programs.

The first (Table XXVII) is the one used perform stepwise regression on the whole set of predictors so as to select out those which the more useful. The reader will note that after the SH data is read in, it is sorted by race which facilitates the subsequent regressions which are performed by race.

The second (Table XXVIII) uses the output from the previous analyses to construct regression models for relevant race/sex groupings. The data is split, for each grouping, into two random samples used as the basis for cross-validation. Predictor and criterion scores, after validation are then output, in raw data form, to mass storage so as to be available for later FORTRAN analysis in estimating cut-offs.

3. FORTRAN Program used to Estimate Cut-offs on Predictors

The raw output produced in the preceding SAS run is used as input data for this program (Table XXIX) along with the appropriate utility values.

The program is written so as to be run on one of the Navy Postgraduate School's terminals which is linked to a Tecktronics plotter, although it could easily be rewritten for a different plotting device. The plotting is performed using subroutines of the School's DISSPLA computer package.

TABLE XXVI

Listing of SAS Set-Up Program: Cohort to File

```
//NESBITTS JOB (2501,0171),'NESBITT ',CLASS=K
//*MAIN ORG=NPGVM1.2501P
// EXEC PGM=IEFBR14
//DD1 DD DISP=(OLD,DELETE),DSN=MSS.S2501.NRATESH
//  EXEC SAS
//SAS.WORK DD SPACE=(CYL,(10,10)),
//             DISP=(NEW,DELETE,DELETE),
//   VOL=SER=(MVS007,MVS012,MVS009,MVS004),UNIT=3350
//FILEIN DD  UNIT=3400-5,VOL=SER=ENLIST,
//   DISP=OLD,DSN=ENLST.ALL.A7678
//FILEOUT DD UNIT=3330V,MSVGP=PUB4B,
//    DISP=(NEW,CATLG),DSN=MSS.S2501.NRATESH,
//    DCB=(BLKSIZE=6400)
//SYSIN DD *
OPTIONS ERRORS=5;
DATA FILEOUT.NRATESH;
     INFILE FILEIN; INPUT
@ 16 ENTRYAGE PIB1. @ 17 RECORDID PIB1.
@ 18 HYEC     PIB1. @ 19 SEX      PIB1.
@ 20 RACE     PIB1. @ 23 MRTLDPND PIB1.
@ 24 TESTFORM PIB1. @ 25 AFQTPCNT PIB1.
@ 26 AFQTGRPS PIB1. @ 27 ASVABGI  PIB1.
@ 28 ASVABNO  PIB1. @ 29 ASVABAD  PIB1.
@ 30 ASVABWK  PIB1. @ 31 ASVABAR  PIB1.
@ 32 ASVABSP  PIB1. @ 33 ASVABMK  PIB1.
@ 34 ASVABEI  PIB1. @ 35 ASVABMC  PIB1.
@ 36 ASVABGS  PIB1. @ 37 ASVABSI  PIB1.
@ 38 ASVABAI  PIB1. @ 39 SERVACCS PIB1.
@ 40 PRIORSRV PIB1. @ 43 ASVABCM  PIB1.
@ 44 ASVABCA  PIB1. @ 45 ASVABCE  PIB1.
@ 46 ASVABCC  PIB1. @ 58 ENTRYYR  PIB1.
@ 61 TERMENLT PIB1. @ 62 ENTRPAYG PIB1.
@ 59 ENTRYMTH PIB1. @ 60 ENTRYDAY PIB1.
@ 65 PROGENLT PIB5. @ 73 BONUSOPT PIB1.
@ 74 ENLSTOPT PIB1. @ 75 YOUTHPRG PIB1.
@ 78 TAPEDATE PIB1. @ 81 TRENLMOS PIB5.
@ 86 TAFMS1   PIB2. @ 88 DPOC1    PIB2.
@ 90 DDOC1    PIB2. @ 92 HYEC1    PIB1.
@ 93 PAYGRDE1 PIB1. @ 94 SERVICE1 PIB1.
@ 95 MRTSTAT1 PIB1. @ 96 NDPNDNT1 PIB1.
@ 97 SPNSPD1  PIB3. @100 ISC1     PIB1.
@101 SEPRT1YR PIB1. @102 SEPRT1MT PIB1.
@103 SEPRT1DY PIB1. @104 BASD1YR  PIB1.
@105 BASD1MTH PIB1. @106 BASD1DAY PIB1.
@107 ETS1YEAR PIB1. @108 ETS1MNTH PIB1.
@109 DOLE1YR  PIB1. @110 DOLE1MTH PIB1.
@113 PEBD1YR  PIB1. @114 PEBD1MTH PIB1.
@115 PEBD1DAY PIB1. @111 CHARSRV1 PIB1.
@112 ELGREUP1 PIB1. @116 FILEFLG1 PIB2.
@118 TAFMS2   PIB2. @120 DPOC2    PIB2.
@122 DDOC2    PIB2. @124 HYEC2    PIB1.
@125 PAYGRDE2 PIB1. @126 SERVICE2 PIB1.
@127 MRTSTAT2 PIB1. @128 NDPNDNT2 PIB1.
@129 SPNSPD2  PIB3. @132 ISC2     PIB1.
@133 SEPRT2YR PIB1. @134 SEPRT2MT PIB1.
@135 SEPRT2DY PIB1. @136 BASD2YR  PIB1.
@137 BASD2MTH PIB1. @138 BASD2DAY PIB1.
@139 ETS2YEAR PIB1. @140 ETS2MNTH PIB1.
@141 DOLE2YR  PIB1. @142 DOLE2MTH PIB1.
@145 PEBD2YR  PIB1. @146 PEBD2MTH PIB1.
@147 PEBD2DAY PIB1. @143 CHARSRV2 PIB1.
@144 ELGREUP2 PIB1. @148 FILEFLG2 PIB2.
@150 TAFMS3   PIB1. @151 TAFMS4   PIB1.
@152 DPOC3    PIB2. @154 DDOC3    PIB2.
```

```
@156 HYEC3     PIB1.  @157 PAYGRDE3  PIB1.
@158 SERVICE3  PIB1.  @159 MRTSTAT3  PIB1.
@160 NDPNDNT3  PIB1.  @161 SPNSPD3   PIB3.
@165 SEPRT3YR  PIB1.  @166 SEPRT3MT  PIB1.
@167 SEPRT3DY  PIB1.  @168 BASD3YR   PIB1.
@169 BASD3MTH  PIB1.  @170 BASD3DAY  PIB1.
@171 ETS3YEAR  PIB1.  @172 ETS3MNTH  PIB1.
@173 DOLE3YR   PIB1.  @174 DOLE3MTH  PIB1.
@177 PEBD3YR   PIB1.  @178 PEBD3MTH  PIB1.
@179 PEBD3DAY  PIB1.  @164 ISC3      PIB1.
@175 CHARSRV3  PIB1.  @176 ELGREUP3  PIB1.
@180 FILEFLG3  PIB2.  @182 FILEMTCH  PIB4.
@186 DOEYRDEP  PIB1.  @187 DOEMTDEP  PIB1.
@188 MNTHSDEP  PIB1.  @189 SPFLGML   PIB1.
@190 DCPGYR    PIB1.  @191 DCPGMNTH  PIB1.
@212 GCT  2. @214 ARI 2. @216 MECH     2.
@218 CLER 2. @220 AFQTS 2. @222 PNEC $4.
@227 CTZNSHIP    $1.  @229 PRIDEPND     $1.
@230 SECDEPND    $1.  @231 BRCL         $2.
@233 GROUPIND    $1.  @234 AUTHRATE     $4.
@240 EDPGYR      $4.  @244 SCHLCODE     $1.
@245 SCHLWVR     $1.  @246 ASTAR        $1.
@247 TSSIND      $1.  @250 PRESRATE     $4.
@254 NUMPG1      $1.  @255 PRRTABRV     $3.
@258 EXAMRATE    $4.  @262 NUMPG2       $1.
@263 EXRTABRV    $3.  @266 TOTLRAW       3.
@269 STDNAVY      2.  @272 PRCODE       $2.
@274 ALTPRCDE    $2.  @276 FINLMULT      5.
@281 FNMLTCUT     5.  @287 PRFFACTR      3.
@290 AWIFACTR     2.  @292 CHNGRATE     $1.
@296 RATEIND     $1.  @297 SPPROIND     $1.
@298 TYPENLST    $2.  @301 MODEST       $1.
@302 NENLSTMT     1.  @303 EAOS   YYMMDD6.
@309 TAS         $4.  @313 OAS          $4.
@317 LOSCODE     $1.  @318 LOSWVR       $1.
@319 SIPG        $4.  @323 TIRWVR       $1.
@324 TIR         $4.  @336 ADBD   YYMMDD6.
@343 EDPG   YYMMDD6.  @349 DTIS          3.
@352 RECFORES     1.  @356 NCHANGES      3.
@384 AGE          2.  @386 NHRCGCT       2.
@388 NHRCAFQT     2.  @390 MENTLGRP     $1.
@391 EDCERTIF    $1.  @392 MOBLDSGN     $1.
@394 HYNDPNDT     2.  @396 GRP4PROG     $2.
@398 SSDUTY      $1.  @399 REGRESRV     $1.
@400 HYPAYGRD    $1.  @401 NOTRCMD      $1.
@402 SSNCHNGE    $1.  @403 TOTPROMO      2.
@405 TOTLDEMO     1.  @406 TOTLAWOL      1.
@407 TOTDESRT     1.  @408 TOTMLTCN      1.
@409 TOTCVLCN     1.  @412 LNGTHSRV     $4.
@416 SCREEN       2.  @418 ATTRITCD     $1.
@419 RECNTC      $1.  @420 RECENLST     $2.
@422 RECPROGM    $1.  @423 RECPRGSC     $2.
@425 RCPGSCRT    $4.  @435 ELSTHIST     $1.
@436 NDAYSE2      4.  @440 NDAYSE3       4.
@444 NDAYSE4      4.  @449 DMDCRATE     $3.
@452 DMDCNEC     $4.  @456 DMDCUIC      $6.
@480 EARNNEC $4. @484 TRAININD $1.
@485 STACTION $1.;
LABEL
ENTRYAGE=AGE OF INDIVIDUAL AT TIME OF ENTRY
RECORDID=RECORD ID--EXAM SCORE, DEP, ACTIVE DUTY
HYEC    =HIGHEST YEAR OF EDUCATION
SEX     =(1) MALE, (2) FEMALE
RACE    =(1) WHITE, (2) BLACK, (3) OTHER
MRTLDPND=MARITAL STATUS/DEPENDENTS
TESTFORM=TEST FORM/ECFA,ASVAB,AFWST,AFQT,OSB...
AFQTPCNT=AFQT PERCENTILE (OR EQUIVALENT)
AFQTGRPS=AFQT GROUPS (5,4C,4B,4A,3B,3A,2,1)
ASVABGI =ASVAB APTITUDE AREA SCORE--SUBSCALE GI
```

```
ASVABNO  =ASVAB APTITUDE AREA SCORE--SUBSCALE NO
ASVABAD  =ASVAB APTITUDE AREA SCORE--SUBSCALE AD
ASVABWK  =ASVAB APTITUDE AREA SCORE--SUBSCALE WK
ASVABAR  =ASVAB APTITUDE AREA SCORE--SUBSCALE AR
ASVABSP  =ASVAB APTITUDE AREA SCORE--SUBSCALE SP
ASVABMK  =ASVAB APTITUDE AREA SCORE--SUBSCALE MK
ASVABEI  =ASVAB APTITUDE AREA SCORE--SUBSCALE EI
ASVABMC  =ASVAB APTITUDE AREA SCORE--SUBSCALE MC
ASVABGS  =ASVAB APTITUDE AREA SCORE--SUBSCALE GS
ASVABSI  =ASVAB APTITUDE AREA SCORE--SUBSCALE SI
ASVABAI  =ASVAB APTITUDE AREA SCORE--SUBSCALE AI
SERVACCS=SERVICE OF ACCESSION (NAVY,2)
PRIORSRV=PRIOR SERVICE (NON-PRIOR SERVICE,1)
ASVABCM  =ASVAB APTITUDE AREA SCORE--SUBSCALE CM
ASVABCA  =ASVAB APTITUDE AREA SCORE--SUBSCALE CA
ASVABCE  =ASVAB APTITUDE AREA SCORE--SUBSCALE CE
ASVABCC  =ASVAB APTITUDE AREA SCORE--SUBSCALE CC
TERMENLT=TERM OF ENLISTMENT (NO. OF YEARS)
ENTRPAYG=ENTRY PAY GRADE (E00--011)
PROGENLT=PROGRAM ENLISTED FOR--SERVICE UNIQUE
BONUSOPT=BONUS OPTION, COMBAT OR NON-COMBAT
ENLSTOPT=ENLISTMENT OPTION
YOUTHPRG=YOUTH & RESERVE TRAINING PROGRAMS
TAPEDATE=MONTH OF FILE ON WHICH RECORD SUBMITTED
TRENLMOS=OCCUP. SPECIAL./RATING CHOICE UPON ENTRY
TAFMS1   =MONTHS OF TOTL. ACTIVE FED. MILIT. SERV.
DPOC1    =D.O.D. PRIMARY OCCUPATION CODE
DDOC1    =D.O.D. DUTY OCCUPATION CODE
HYEC1    =HIGHEST YEAR OF EDUCATION
PAYGRDE1=PAY GRADE AS-OF-DATE-OF-FILE/SEPARATION
SERVICE1=SERVICE CODE (2, NAVY)
MRTSTAT1=MARITAL STATUS (1,OTHER, 2,MARRIED)
NDPNDNT1=NUMBER OF DEPENDENTS (1, NONE)
SPNSPD1  =SEPARATION PROGRAM DESIGNATOR
ISC1     =INTER-SERVICE SEPARATION CODE
SEPRT1YR=YEAR OF SEPARATION (2ND DMDC SECTION)
SEPRT1MT=MONTH OF SEPARATION (2ND DMDC SECTION)
SEPRT1DY=DAY OF SEPARATION (2ND DMDC SECTION)
BASD1YR  =YEAR OF ACTIVE DUTY BASE DATE
BASD1MTH=MONTH OF ACTIVE DUTY BASE DATE
BASD1DAY=DAY OF ACTIVE DUTY BASE DATE
ETS1YEAR=ESTIMATED YEAR OF FULFILLED ACTIVE DUTY
ETS1MNTH=ESTIMATED MONTH OF FULFILLED ACTIVE DUTY
CHARSRV1=CHARACTER OF SERVICE
ELGREUP1=REENLISTMENT ELIGIBILITY
PEBD1YR  =YEAR OF PAY ENTRY BASE DATE
PEBD1MTH=MONTH OF PAY ENTRY BASE DATE
PEBD1DAY=DAY OF PAY ENTRY BASE DATE
ENTRYYR  =YEAR OF ENTRY TO ACTIVE/D.E.P.
ENTRYMTH=MONTH OF ENTRY TO ACTIVE/D.E.P.
ENTRYDAY=DAY OF ENTRY TO ACTIVE/D.E.P.
SEPRT1YR=YEAR OF SEPARATION (2ND DMDC SECTION)
SEPRT1MT=MONTH OF SEPARATION (2ND DMDC SECTION)
SEPRT1DY=DAY OF SEPARATION (2ND DMDC SECTION)
BASD1YR  =YEAR OF ACTIVE DUTY BASE DATE
BASD1MTH=MONTH OF ACTIVE DUTY BASE DATE
BASD1DAY=DAY OF ACTIVE DUTY BASE DATE
ETS1YEAR=ESTIMATED YEAR OF FULFILLED ACTIVE DUTY
ETS1MNTH=ESTIMATED MONTH OF FULFILLED ACTIVE DUTY
PEBD1YR  =YEAR OF PAY ENTRY BASE DATE
PEBD1MTH=MONTH OF PAY ENTRY BASE DATE
PEBD1DAY=DAY OF PAY ENTRY BASE DATE
FILEFLG1=FILE FLAG NO. 1
PEBD2YR  =YEAR OF PAY ENTRY BASE DATE
PEBD2MTH=MONTH OF PAY ENTRY BASE DATE
PEBD2DAY=DAY OF PAY ENTRY BASE DATE
SEPRT2YR=YEAR OF SEPARATION (3RD DMDC SECTION)
SEPRT2MT=MONTH OF SEPARATION (3RD DMDC SECTION)
SEPRT2DY=DAY OF SEPARATION (3RD DMDC SECTION)
```

```
BASD2YR =YEAR OF ACTIVE DUTY BASE DATE
BASD2MTH=MONTH OF ACTIVE DUTY BASE DATE
BASD2DAY=DAY OF ACTIVE DUTY BASE DATE
ETS2YEAR=ESTIMATED YEAR OF FULFILLED ACTIVE DUTY
ETS2MNTH=ESTIMATED MONTH OF FULFILLED ACTIVE DUTY
PEBD2YR =YEAR OF PAY ENTRY BASE DATE
PEBD2MTH=MONTH OF PAY ENTRY BASE DATE
PEBD2DAY=DAY OF PAY ENTRY BASE DATE
TAFMS2  =MONTHS OF TOTL. ACTIVE FED. MILIT. SERV.
DPOC2   =D.O.D. PRIMARY OCCUPATION CODE
DDOC2   =D.O.D. DUTY OCCUPATION CODE
HYEC2   =HIGHEST YEAR OF EDUCATION
PAYGRDE2=PAY GRADE AS-OF-DATE-OF-FILE/SEPARATION
SERVICE2=SERVICE CODE (2, NAVY)
MRISTAT2=MARITAL STATUS (1,OTHER, 2,MARRIED)
NDPNDNT2=NUMBER OF DEPENDENTS (1, NONE)
SPNSPD2 =SEPARATION PROGRAM DESIGNATOR
ISC2    =INTER-SERVICE SEPARATION CODE
CHARSRV2=CHARACTER OF SERVICE
ELGREUP2=REENLISTMENT ELIGIBILITY
FILEFLG2=FILE FLAG NO. 2
PEBD3YR =YEAR OF PAY ENTRY BASE DATE
PEBD3MTH=MONTH OF PAY ENTRY BASE DATE
PEBD3DAY=DAY OF PAY ENTRY BASE DATE
SEPRT3YR=YEAR OF SEPARATION (4TH DMDC SECTION)
SEPRT3MT=MONTH OF SEPARATION (4TH DMDC SECTION)
SEPRT3DY=DAY OF SEPARATION (4TH DMDC SECTION)
BASD3YR =YEAR OF ACTIVE DUTY BASE DATE
BASD3MTH=MONTH OF ACTIVE DUTY BASE DATE
BASD3DAY=DAY OF ACTIVE DUTY BASE DATE
ETS3YEAR=ESTIMATED YEAR OF FULFILLED ACTIVE DUTY
ETS3MNTH=ESTIMATED MONTH OF FULFILLED ACTIVE DUTY
PEBD3YR =YEAR OF PAY ENTRY BASE DATE
PEBD3MTH=MONTH OF PAY ENTRY BASE DATE
PEBD3DAY=DAY OF PAY ENTRY BASE DATE
TAFMS3  =MONTHS OF TOTL. ACTIVE FED. MILIT. SERV.
TAFMS4  =MONTHS OF TOTL. ACTIVE FED. MILIT. SERV.
DPOC3   =D.O.D. PRIMARY OCCUPATION CODE
DDOC3   =D.O.D. DUTY OCCUPATION CODE
HYEC3   =HIGHEST YEAR OF EDUCATION
PAYGRDE3=PAY GRADE AS-OF-DATE-OF-FILE/SEPARATION
SERVICE3=SERVICE CODE (2, NAVY)
MRISTAT3=MARITAL STATUS (1,OTHER, 2,MARRIED)
NDPNDNT3=NUMBER OF DEPENDENTS (1, NONE)
SPNSPD3 =SEPARATION PROGRAM DESIGNATOR
ISC3    =INTER-SERVICE SEPARATION CODE
CHARSRV3=CHARACTER OF SERVICE
ELGREUP3=REENLISTMENT ELIGIBILITY
FILEFLG3=FILE FLAG NO. 2
FILEMTCH=4-BYTE BINARY FILE MATCH INDICATORS
DOEYRDEP=DOE YEAR INTO D.E.P.
DOEMTDEP=DOE MONTH INTO D.E.P.
MNTHSDEP=MONTHS IN D.E.P.
SPFLGML =SPANISH FLAG MASTER/LOSS
DCPGMNTH=MONTH OF DCPG
DCPGYR  =YEAR OF DCPG
GCT     =BASIC BATTERY GCT
ARI     =BASIC BATTERY ARI
MECH    =BASIC BATTERY MECH
CLER    =BASIC BATTERY CLER
ENEC    =NAVY ENLISTED JOB CODE
CTZNSHIP=CITIZENSHIP CODE
BRCL    =BRANCH/CLASS
GROUPIND=GROUP INDICATOR
AUTHRATE=AUTHORIZED RATE (ABBR.)
EDPGYR  =EFFECTIVE DATE OF PAY GRADE
SCHLCODE=SCHOOL CODE
SCHLWVR =SCHOOL WAIVER
PRESRATE=PRESENT RATE CODE
```

```
PRRTABRV=PRESENT RATE (ABBR.)
EXAMRATE=EXAMINATION RATE CODE
EXRTABRV=EXAMINATION RATE (ABBR.)
TOTLRAW =TOTAL RAW SCORE
STDNAVY =STANDARDIZED NAVY SCORE
PRCODE  =PROCESS CODE
ALTPRCDE=ALTERNATE PROCESS CODE
FINLMULT=CANDIDATE''S FINAL MULTIPLE
FNMLTCUT=FINAL MULTIPLE CUT
PRFFACTR=PERFORMANCE FACTOR
AWIFACTR=AWI FACTOR
CHNGRATE=CHANGE OF RATE INDICATOR
NENLSTMT=NUMBER OF ENLISTMENTS
EAOS     =EXPIRATION OF ACTIVE OBLIGATED SERVICE
TAS      =TOTAL ACTIVE SERVICE
CAS      =OTHER ACTIVE SERVICE
SIPG     =SERVICE IN PAY GRADE
LOSCODE =LENGTH OF SERVICE
LOSWVR  =LENGTH OF SERVICE WAIVER
TIRWVR  =TIME IN RATE WAIVER
TIR      =TIME IN RATE
ADBD     =ACTIVE DUTY BASE DATE
EDPG     =EFFECTIVE DATE OF PAY GRADE
DTIS     =DRILL TIME IN SERVICE
NCHANGES=NUMBER OF CHANGES/ENTRIES IN NHRC FILE
AGE      =CANDIDATE''S CURRENT AGE
NHRCGCT =NHRC FILE''S GENRL. CLASSIFICATION TEST
NHRCAFQT=NHRC FILE''S ARMED FORCES QUALIFY. TEST
MENTLGRP=MENTAL GROUP CODE
EDCERTIF=EDUCATION CERTIFICATE
MOBLDSGN=MILITARY OBLIGATION DESIGNATOR
HYNDPNDT=HIGHEST NUMBER OF PRIMARY DEPENDENTS
GRP4PROG=GROUP IV (100K) PROGRAM CODE
SSDUTY  =SEA-SHORE DUTY INDICATOR
REGRESRV=REGULAR RESERVE INDICATOR
HYPAYGRD=HIGHEST PAY GRADE
NOTRCMD =NOT RECOMMENDED FOR RE-ENLISTMENT
SSNCHNGE=SOCIAL SECURITY/NAME CHANGE
TOTPROMO=TOTAL PROMOTIONS
TOTLDEMO=TOTAL DEMOTIONS
TOTLAWOL=TOTAL UA/AWOL
TOTDESRT=TOTAL DESERTIONS
TOTMLTCN=TOTAL MILITARY CONFINEMENTS
TOTCVLCN=TOTAL CIVILIAN CONFINEMENTS
LNGTHSRV=LENGTH OF SERVICE
SCREEN  =SCREEN SCORE
ATTRITCD=ATTRITION INDICATOR
RECNTC  =RECRUIT NAVAL TRAINING COMMAND
RECENLST=RECRUIT TYPE ENLISTMENT
RECPRCGM=RECRUIT PROGRAM AT ENLISTMENT
RECPRGSC=RECRUIT PROGRAM/SCHOOL
RCPGSCRT=RECRUIT PROGRAM/SCHOOL RATE
ELSTHIST=ENLISTED HISTORY STATUS
NDAYSE2 =COMPUTED NUMBER OF DAYS TO E-2 RATING
NDAYSE3 =COMPUTED NUMBER OF DAYS TO E-3 RATING
NDAYSE4 =COMPUTED NUMBER OF DAYS TO E-4 RATING
DCLE1YR =YEAR OF LATEST RE-ENLISTMENT
DCLE1MTH=MONTH OF LATEST RE-ENLISTMENT
DCLE2YR =YEAR OF LATEST RE-ENLISTMENT
DCLE2MTH=MONTH OF LATEST RE-ENLISTMENT
DCLE3YR =YEAR OF LATEST RE-ENLISTMENT
DCLE3MTH=MONTH OF LATEST RE-ENLISTMENT
DMDCRATE=FINAL RATING AS LISTED BY D.M.D.C.
DMDCNEC =FINAL N.E.C. AS LISTED BY D.M.D.C.
DMDCUIC =FINAL U.I.C. AS LISTED BY D.M.D.C.
CONVDATE=CONVENING DATE FOR NITRAS COURSE
GRADDATE=GRADUATION DATE FOR NITRAS COURSE
TRANDATE=TRANSACTION DATE FOR NITRAS RECORD
EARNNEC =DID CANDIDATE EARN AN NEC?
```

```
TRAININD=TRAINING INDICATOR
STACTION=STUDENT ACTION CODES (PASS, P, ETC.);
IF DMDCRATE='SH' OR PRRTABRV='SH' OR
             RCPGSCRT='2490' OR EXAMRATE='2490';
PROC FREQ;
  TABLES DMDCRATE PRRTABRV RCPGSCRT
             AUTHRATE PRESRATE EXAMRATE;
  TITLE ATTEMPT AT FORMING AN SH FILE;
/*
//
```

TABLE XXVII

Program Listing of Stepwise Regression for SH

```
//NESBITTS JOE (2501,0171),'NESBITT ',CLASS=C
//*MAIN LINES=(25),ORG=NPGVM1.2501P
//  EXEC SAS
//SAS.WORK DD SPACE=(CYL,(25,50)),
//         DISP=(NEW,DELETE,DELETE)
//FILEIN DD  DISP=SHR,DSN=MSS.S2501.NRATESH
//SYSIN DD *
OPTIONS LS=80 NOCENTER NODATE;
DATA;SET FILEIN.NRATESH;
* -------------------------------------------------
| NUMBER OF YEARS OF EDUCATION IS CONVERTED FROM |
| ITS DMDC ORDINAL CODING (1-13) TO A numeric.   |
|-------------------------------------------------;
IF HYEC=1 THEN CHYEC=3.5;IF HYEC=2 THEN CHYEC=8;
IF HYEC=3 THEN CHYEC=9; IF HYEC=4 THEN CHYEC=10;
IF HYEC=5 THEN CHYEC=11;IF HYEC=6 THEN CHYEC=12;
IF HYEC=7 THEN CHYEC=13;IF HYEC=8 THEN CHYEC=14;
IF HYEC=9 THEN CHYEC=15;IF HYEC=10 THEN CHYEC=16;
IF HYEC=11 THEN CHYEC=18;IF HYEC=12 THEN CHYEC=20;
IF HYEC=13 THEN CHYEC=11.5;
* ----------------------------------------
|  VALIDITY VALUE SCREENS                 |
|----------------------------------------;
IF ENTRYAGE GE 35 THEN ENTRYAGE=35;
IF ENTRYAGE >= 17;IF TOTPROMO<=5;
IF SCHLCODE='A' THEN SCHLCODE='1';
ELSE SCHLCODE='0';NUSCHCDE=SCHLCODE+0;
NUATTRIT=ATTRITCD+0; IF NUATTRIT=2 THEN
NUATTRIT=1;ELSE NUATTRIT=0;
NUNOTRC=NOTRCMD+0;
NUHYPAY=HYPAYGRD+0;
YEAR=SUBSTR(LNGTHSRV,1,2);
MONTH=SUBSTR(LNGTHSRV,3,2);
YEARS=YEAR+0;MONTHS=MONTH+0;
LOSMNTHS=YEARS*12+MONTHS;
IF MONTHS >= 6 THEN YEARS=YEARS+1;
ONEYEAR=0;IF LOSMNTHS >= 12 THEN ONEYEAR=1;
TWOYEAR=0;IF LOSMNTHS >= 24 THEN TWOYEAR=1;
THRYEAR=0;IF LOSMNTHS >= 36 THEN THRYEAR=1;
FORYEAR=0;IF LOSMNTHS >= 48 THEN FORYEAR=1;
LABEL LOSMNTHS=TOTAL LOS IN MONTHS (NUMERIC);
* ------------------------------------
| the highest paygrade is determined |
|------------------------------------;
IF FILEFLG1=8209 THEN PAYGRADE=PAYGRDE1;
IF FILEFLG1 NE 8209 THEN PAYGRADE=PAYGRDE3;
IF PAYGRADE>=2 THEN NDAYSE22=NDAYSE2;ELSE NDAYSE22=0;
IF PAYGRADE>=3 THEN NDAYSE23=NDAYSE3;ELSE NDAYSE23=0;
IF PAYGRADE>=4 THEN NDAYSE24=NDAYSE4;ELSE NDAYSE24=0;
IF NDAYSE22=0 THEN NDAYSE22=1275;
IF NDAYSE23=0 THEN NDAYSE23=1461;
IF NDAYSE24=0 THEN NDAYSE24=1750;
 TIMEE2=0.;
 TIMEE3=0.;
 TIMEE4=0.;
 TIMEE2=(ROUND(((NDAYSE22)/(365/12))*10)/10);
 TIMEE3=(ROUND(((NDAYSE23)/(365/12))*10)/10);
 TIMEE4=(ROUND(((NDAYSE24)/(365/12))*10)/10);
LABEL
 TIMEE2 =ROUNDED NUMBER OF MONTHS TO E2
 TIMEE3 =ROUNDED NUMBER OF MONTHS TO E3
 TIMEE4 =ROUNDED NUMBER OF MONTHS TO E4;
* ------------------------------------
| DEFINES "ELIGIBILITY TO RE-ENLIST".|
```

```
-------------------------------------;
IF FILEFLG1=8209 THEN ELIGREUP=1;
IF ((FILEFLG1 NE 8209) AND (ISC3 GT 0) AND
(ELGREUP3 EQ 1)) THEN ELIGREUP=1;ELSE ELIGREUP=0;
* ---------------------------------------
| defines "ACHIEVED E-4", IN JOINT      | |
| CONSIDERATION OF THE D.M.D.C. AND N.H.R.C. |
-----------------------------------------;
IF ((PAYGRADE GE 4) AND (HYPAYGRD GE 4)) THEN ACHVDE4=1;
IF ((PAYGRADE LT 4) OR (HYPAYGRD LT 4)) THEN ACHVDE4=0;
* ------------------------------------
| defines "RATED" VERSUS "NOT-RATED".|
| -----------------------------------;
IF ((DMDCRATE NE '.') AND (DMDCRATE NE ' ')
AND (SERVACCS EQ 2) AND (SERVICE1 EQ 2) AND
((PAYGRADE GE 4) AND (HYPAYGRD GE 4))) THEN
 RATED=1;ELSE RATED=0;
* -------------------------------------------
| eliminates invalid asvab subscale scores |
-------------------------------------------;
IF ASVABGI<=15;IF ASVABNO<=50;IF ASVABAD<=30;
IF ASVABWK<=30;IF ASVABAR<=20;IF ASVABSP<=20;
IF ASVABMK<=20;IF ASVABEI<=30;IF ASVABMC<=30;
IF ASVABGS<=20;IF ASVABSI<=20;IF ASVABAI<=20;
IF ASVABCM<=30;IF ASVABCA<=20;IF ASVABCE<=30;
IF ASVABCC<=30;
* ---------------------------------------------
| establishes entry groups:                    |
|             entry           exam        dmdc |
|  (1)         yes            yes         yes  |
|  (2)         yes            yes         no   |
|  (3)         yes            no          yes  |
|  (4)         yes            no          no   |
|  (5)         no             yes         yes  |
|  (6)         no             yes         no   |
|  (7)         no             no          yes  |
|----------------------------------------------;
IF (RCPGSCRT='2490' AND EXAMRATE='2490'
AND DMDCRATE='SH') THEN ENTRYGRP=1;
IF (RCPGSCRT='2490' AND EXAMRATE='2490'
AND DMDCRATE NE 'SH') THEN ENTRYGRP=2;
IF (RCPGSCRT='2490' AND EXAMRATE NE '2490'
 AND DMDCRATE='SH') THEN ENTRYGRP=3;
IF (RCPGSCRT='2490' AND EXAMRATE NE '2490'
 AND DMDCRATE NE 'SH') THEN ENTRYGRP=4;
IF (RCPGSCRT NE '2490' AND EXAMRATE='2490'
AND DMDCRATE='SH') THEN ENTRYGRP=5;
IF (RCPGSCRT NE '2490' AND EXAMRATE='2490'
 AND DMDCRATE NE 'SH') THEN ENTRYGRP=6;
IF (RCPGSCRT NE '2490' AND EXAMRATE NE '2490'
 AND DMDCRATE='SH') THEN ENTRYGRP=7;
* -------------------------
|  NEW VARIABLE "DEPNDNTS" |
|-------------------------;
IF MRTLDPND=10 THEN DEPNDNTS=0;
IF MRTLDPND GT 10 THEN DEPNDNTS=1;

GOODGUY=20;
IF LOSMNTHS < 48 THEN GOODGUY =10;
IF LOSMNTHS >= 48 AND RATED=1
AND ELIGREUP=1 THEN GOODGUY =30;
BADGUY=30;
IF LOSMNTHS < 48 AND (ISC3 <=87 AND
ISC3>=60)  THEN BADGUY =10;
IF LOSMNTHS < 48 AND TOTLAWOL > 0  THEN BADGUY =10;
IF LOSMNTHS < 48 AND TOTDESRT > 0  THEN BADGUY =10;
IF LOSMNTHS < 48 AND TOTMLTCN > 0  THEN BADGUY =10;
IF LOSMNTHS < 48 AND TOTCVLCN > 0  THEN BADGUY =10;
IF LOSMNTHS < 48 AND BADGUY=30  THEN BADGUY =20;
```

```
IF LOSMNTHS >= 48 AND TOTLDEMO > 0 THEN BADGUY =20;
IF LOSMNTHS >= 48 AND ELIGREUP = 0 THEN BADGUY =20;
LABEL
DEPNDNTS=SINGLE, NO KIDS 0/OTHERS 1
CHYEC    =CONVERTED NUMBER OF YEARS OF EDUCATION
NUHYPAY  =NHRC FILE--HIGHEST PAYGRADE ATTAINED
NUSCHCDE=ADVANCEMENT FILE--''A'' SCHOOL COMPLETED
NUATTRIT=NHRC FILE--ATTRITION CODES
NUNCTRC  =NHRC--NOT RECOMMENDED FOR RE-ENLISTMENT
NDAYSE22=SCREENED NUMBER OF DAYS TO E-2
NDAYSE23=SCREENED NUMBER OF DAYS TO E-3
NDAYSE24=SCREENED NUMBER OF DAYS TO E-4
ELIGREUP=ELIGIBLE TO RE-ENLIST
ATTRITC2=DMDC-BASED STANDARD ATTRITION MEASURE
ATTRITC3=DMDC-BASED NEGATIVE ATTRITION MEASURE
PAYGRADE=DMDC-BASED HIGHEST PAY-GRADE ATTAINED
ACHVDE4  =DMDC & NHRC CONCORDANT E-4 ACHIEVED
RATED    =ACCESSED & MOST RECENTLY NAVY--MADE E-4
ENTRYGRP=7 ENTRY GROUPS, NHRC/EXAMRATE/DMDCRATE;
IF ENTRYGRP NE 4;
    RANDNUM1=ROUND(2*UNIFORM(12354));
 IF RANDNUM1=0 THEN RANDNUM1=2;
IF BASD1YR <= 78 OR (BASD1YR = 78 AND BASD1MTH <= 9);
IF ENTRYGRP NE 4;
IF ISC1 > 16 OR ISC1 < 10;*  MEDICAL DISCHARGES;
IF ISC2 > 16 OR ISC2 < 10;*  MEDICAL DISCHARGES;
IF ISC3 > 16 OR ISC3 < 10;*  MEDICAL DISCHARGES;
IF ISC1 > 33 OR ISC1 < 30;*  DEATHS  DISCHARGEES;
IF ISC2 > 33 OR ISC2 < 30;*  DEATHS  DISCHARGEES;
IF ISC3 > 33 OR ISC3 < 30;*  DEATHS  DISCHARGEES;
IF ISC1 > 42 OR ISC1 < 40;*  OFFICER ENTRANTS;
IF ISC2 > 42 OR ISC2 < 40;*  OFFICER ENTRANTS;
IF ISC3 > 42 OR ISC3 < 40;*  OFFICER ENTRANTS;
IF SEX=1;* SELECTS ONLY MALES;
 PROC SORT;
   BY RACE;
PROC STEPWISE;
 BY RACE;
 MODEL LOSMNTHS GOODGUY BADGUY= AFQTPCNT
   AFQTGRPS ASVABGI ASVABNO ASVABAD   ASVABWK
ASVABAR ASVABSP ASVABMK ASVABEI ASVABMC ASVABGS
   ASVABSI ASVABAI ASVABCM ASVABCA ASVABCE ASVABCC
   ENTRYAGE CHYEC ENTRPAYG MRTSTAT1
   SCREEN DEPNDNTS/STEPWISE;
   TITLE STEPWISE SELECTION OF PREDICTORS
   BY RACE SH SAMPLE;
//
```

TABLE XXVIII

Program Listing to Output & Cross Validate

```
//NESBITTS JOB (2501,C171),'NESBITT ',CLASS=C
//*MAIN LINES=(25),ORG=NPGVM1.2501P
//  EXEC SAS
//SAS.WORK DD SPACE=(CYL,(25,50)),
//          DISP=(NEW,DELETE,DELETE)
//FILEIN DD  DISP=SHR,DSN=MSS.S2501.NRATESH
//FT31F001 DD UNIT=3330V,MSVGP=PUB4Z,
//     DISP=(NEW,CATLG),DSN=MSS.S2501.SHLMN,
//        DCB=(RECFM=FB,LRECL=80,BLKSIZE=3120)
//FT32F001 DD UNIT=3330V,MSVGP=PUB4Z,
//     DISP=(NEW,CATLG),DSN=MSS.S2501.SHLWH,
//        DCB=(RECFM=FB,LRECL=80,BLKSIZE=3120)
//FT33F001 DD UNIT=3330V,MSVGP=PUB4Z,
//     DISP=(NEW,CATLG),DSN=MSS.S2501.SHLBL,
//        DCB=(RECFM=FB,LRECL=80,BLKSIZE=3120)
//FT34F001 DD UNIT=3330V,MSVGP=PUB4Z,
//     DISP=(NEW,CATLG),DSN=MSS.S2501.SHLOT,
//        DCB=(RECFM=FB,LRECL=80,BLKSIZE=3120)
//SYSIN DD *
data;set filein.nratesh;
ONEYEAR=0;IF LOSMNTHS >= 12 THEN ONEYEAR=1;
TWOYEAR=0;IF LOSMNTHS >= 24 THEN TWOYEAR=1;
THRYEAR=0;IF LOSMNTHS >= 36 THEN THRYEAR=1;
FORYEAR=0;IF LOSMNTHS >= 48 THEN FORYEAR=1;
IF SEX=1;* SELECTS ONLY MALES;
DATA TRIALDAT; SET DATA1;
IF RACE>=0;* SELECTS ALL RACES;
DATA DIFFRNTZ; SET TRIALDAT;
DATA DERIV8;SET DIFFRNTZ;IF RANDNUM1>1;
DATA VALID8;SET DIFFRNTZ;IF RANDNUM1=1;
PROC REG DATA=DERIV8 SIMPLE OUTEST=B01;LOSM1HAT:MODEL
 LOSMNTHS=ENTRYAGE ASVABGI ASVABAI ASVABCA
 SCREEN MRTSTAT1 ENTRPAYG ASVABGS ASVABAD
 ASVABCC ASVABMC ASVABAR AFQTGRPS DEPNDNTS/STB;
TITLE MODEL ALL SHS FOR TOTAL LOS SAMPLE ONE;
PROC REG DATA=VALID8 SIMPLE OUTEST=B02;LOSM2HAT:MODEL
 LOSMNTHS=ENTRYAGE ASVABGI ASVABAI ASVABCA
 SCREEN MRTSTAT1 ENTRPAYG ASVABGS ASVABAD
 ASVABCC ASVABMC ASVABAR AFQTGRPS DEPNDNTS/STB;
TITLE MODEL ALL SHS FOR TOTAL LOS SAMPLE TWO;
PROC SCORE OUT=B01PRED TYPE=OLS SCORE=B01
DATA=VALID8 PREDICT ; VAR ENTRYAGE ASVABGI ASVABAI
 ASVABCA SCREEN MRTSTAT1 ENTRPAYG ASVABGS ASVABAD
 ASVABCC ASVABMC ASVABAR AFQTGRPS DEPNDNTS;
PROC SCORE OUT=B02PRED TYPE=OLS SCORE=B02
DATA=DERIV8 PREDICT ; VAR ENTRYAGE ASVABGI ASVABAI
 ASVABCA SCREEN MRTSTAT1 ENTRPAYG ASVABGS ASVABAD
 ASVABCC ASVABMC ASVABAR AFQTGRPS DEPNDNTS;
PROC CORR DATA=B01PRED;VAR
          LOSMNTHS LOSM1HAT;
TITLE  ALL SHS  FIRST VALIDITY COEFFICIENT;
PROC CORR DATA=B02PRED;VAR
          LOSMNTHS LOSM2HAT;
TITLE  ALL SHS  SECOND VALIDITY COEFFICIENT;
PROC SCORE OUT=ALLPRED TYPE=OLS SCORE=B01
DATA=DIFFRNTZ PREDICT; VAR ENTRYAGE ASVABGI ASVABAI
 ASVABCA SCREEN MRTSTAT1 ENTRPAYG ASVABGS ASVABAD
 ASVABCC ASVABMC ASVABAR AFQTGRPS DEPNDNTS;
PROC SCORE OUT=TWOPRED TYPE=OLS SCORE=B02
DATA=allpred PREDICT; VAR ENTRYAGE ASVABGI ASVABAI
 ASVABCA SCREEN MRTSTAT1 ENTRPAYG ASVABGS ASVABAD
 ASVABCC ASVABMC ASVABAR AFQTGRPS DEPNDNTS;
DATA TABLE;SET TWOPRED;
PRDICTOR=(LOSM1HAT+LOSM2HAT)/2;
```

```
PROC SORT;        BY PRDICTOR;
TITLE  SORTING BY PRDICTOR FOR ALL SH MEN;
DATA STORE;SET TABLE;FILE FT31F001;
PUT (FORYEAR PRDICTOR THRYEAR RACE) (10.7);
DATA TRIALDAT; SET DATA1;
IF RACE=1;* SELECTS WHITES SHS;
DATA DIFFRNTZ; SET TRIALDAT;
DATA DERIV8;SET DIFFRNTZ;IF RANDNUM1>1;
DATA VALID8;SET DIFFRNTZ;IF RANDNUM1=1;
PROC REG DATA=DERIV8 SIMPLE OUTEST=B01;LOSM1HAT:MODEL
 LOSMNTHS=ENTRPAYG ASVABGI ASVABAI ASVABMC SCREEN
 MRTSTAT1 AFQTGRPS DEPNDNTS/ STB;
TITLE MODEL WHITE SHS FOR LOS SAMPLE ONE;
PROC REG DATA=VALID8 SIMPLE OUTEST=B02;LOSM2HAT:MODEL
 LOSMNTHS=ENTRPAYG ASVABGI ASVABAI ASVABMC SCREEN
 MRTSTAT1 AFQTGRPS DEPNDNTS/ STB;
TITLE MODEL WHITE SHS FOR LOS SAMPLE TWO;
PROC SCORE OUT=B01PRED TYPE=OLS SCORE=B01
DATA=VALID8 PREDICT ; VAR   ENTRPAYG ASVABGI ASVABAI
ASVABMC SCREEN MRTSTAT1 AFQTGRPS DEPNDNTS;
PROC SCORE OUT=B02PRED TYPE=OLS SCORE=B02
DATA=DERIV8 PREDICT ; VAR   ENTRPAYG ASVABGI ASVABAI
ASVABMC SCREEN MRTSTAT1 AFQTGRPS DEPNDNTS;
PROC CORR DATA=B01PRED;VAR
         LOSMNTHS LOSM1HAT;
TITLE WHITES    FIRST VALIDITY COEFFICIENT;
PROC CORR DATA=B02PRED;VAR
         LOSMNTHS LOSM2HAT;
TITLE WHITES    SECOND VALIDITY COEFFICIENT;
PROC SCORE OUT=ALLPRED TYPE=OLS SCORE=B01
DATA=DIFFRNTZ PREDICT ; VAR   ENTRPAYG ASVABGI ASVABAI
ASVABMC SCREEN MRTSTAT1 AFQTGRPS DEPNDNTS;
PROC SCORE OUT=TWOPRED TYPE=OLS SCORE=B02
DATA=allpred PREDICT ; VAR   ENTRPAYG ASVABGI ASVABAI
ASVABMC SCREEN MRTSTAT1 AFQTGRPS DEPNDNTS;
DATA TABLE;SET TWOPRED;
PRDICTOR=(LOSM1HAT+LOSM2HAT)/2;
PROC SORT;       BY PRDICTOR;
TITLE  SORTING BY PRDICTOR FOR WHITE  SH MEN;
DATA STORE;SET TABLE;FILE FT32F001;
PUT (FORYEAR PRDICTOR THRYEAR RACE) (10.7);
DATA TRIALDAT; SET DATA1;
IF RACE=2;* SELECTS BLACKS    ;
DATA DIFFRNTZ; SET TRIALDAT;
DATA DERIV8;SET DIFFRNTZ;IF RANDNUM1>1;
DATA VALID8;SET DIFFRNTZ;IF RANDNUM1=1;
PROC REG DATA=DERIV8 SIMPLE OUTEST=B01;LOSM1HAT:MODEL
 LOSMNTHS=ASVABCA SCREEN MRTSTAT1 ASVABCC AFQTGRPS/STB;
TITLE MODEL SH BLACKS FOR LOS  SAMPLE ONE;
PROC REG DATA=VALID8 SIMPLE OUTEST=B02;LOSM2HAT:MODEL
 LOSMNTHS=ASVABCA SCREEN MRTSTAT1 ASVABCC AFQTGRPS/STB;
TITLE MODEL SH BLACKS FOR LOS  SAMPLE TWO;
PROC SCORE OUT=B01PRED TYPE=OLS SCORE=B01
DATA=VALID8 PREDICT; VAR ASVABCA SCREEN
MRTSTAT1 ASVABCC AFQTGRPS;
PROC SCORE OUT=B02PRED TYPE=OLS SCORE=B02
DATA=DERIV8 PREDICT; VAR ASVABCA SCREEN
MRTSTAT1 ASVABCC AFQTGRPS;
PROC CORR DATA=B01PRED;VAR
         LOSMNTHS LOSM1HAT;
TITLE  BLACKS  FIRST VALIDITY COEFFICIENT;
PROC CORR DATA=B02PRED;VAR
         LOSMNTHS LOSM2HAT;
TITLE  BLACKS  SECOND VALIDITY COEFFICIENT;
PROC SCORE OUT=ALLPRED TYPE=OLS SCORE=B01
DATA=DIFFRNTZ PREDICT; VAR ASVABCA SCREEN
MRTSTAT1 ASVABCC AFQTGRPS;
PROC SCORE OUT=TWOPRED TYPE=OLS SCORE=B02
DATA=ALLPRED PREDICT; VAR ASVABCA SCREEN
```

```
MRISTAT1 ASVABCC AFQTGRPS;
DATA TABLE;SET TWOPRED;
PRDICTOR=(LOSM1HAT+LOSM2HAT)/2;
PROC SORT;        BY PRDICTOR;
TITLE  SORTING BY PRDICTOR FOR BLACK SH MEN;
DATA STORE;SET TABLE;FILE FT33F001;
PUT (FORYEAR PRDICTOR THRYEAR RACE) (10.7);
DATA TRIALDAT; SET DATA1;
IF RACE=3;* SELECTS OTHERS SHS;
DATA DIFFRNTZ; SET TRIALDAT;
DATA DERIV8;SET DIFFRNTZ;IF RANDNUM1>1;
DATA VALID8;SET DIFFRNTZ;IF RANDNUM1=1;
PROC REG DATA=DERIV8 SIMPLE OUTEST=B01;LOSM1HAT:MODEL
 LOSMNTHS=ASVABGI ASVABAR ASVABGS ASVABMC ENTRYAGE
    ASVABAD/ STB;
TITLE MODEL OTHERS SHS FOR LOS SAMPLE ONE;
PROC REG DATA=VALID8 SIMPLE OUTEST=B01;LOSM2HAT:MODEL
 LOSMNTHS=ASVABGI ASVABAR ASVABGS ASVABMC ENTRYAGE
    ASVABAD/ STB;
TITLE MODEL OTHERS SHS FOR LOS SAMPLE TWO;
PROC SCORE OUT=B01PRED TYPE=OLS SCORE=B01
DATA=VALID8 PREDICT ; VAR    ASVABGI ASVABAR
ASVABGS ASVABMC ENTRYAGE ASVABAD;
PROC SCORE OUT=B02PRED TYPE=OLS SCORE=B02
DATA=DERIV8 PREDICT ; VAR    ASVABGI ASVABAR
ASVABGS ASVABMC ENTRYAGE ASVABAD;
PROC CORR DATA=B01PRED;VAR
          LOSMNTHS LOSM1HAT;
TITLE  OTHERS    FIRST VALIDITY COEFFICIENT;
PROC CORR DATA=B02PRED;VAR
          LOSMNTHS LOSM2HAT;
TITLE  OTHERS    SECOND VALIDITY COEFFICIENT;
PROC SCORE OUT=ALLPRED TYPE=OLS SCORE=B01
DATA=DIFFRNTZ PREDICT ; VAR    ASVABGI ASVABAR
ASVABGS ASVABMC ENTRYAGE ASVABAD;
PROC SCORE OUT=TWOPRED TYPE=OLS SCORE=B02
DATA=ALLPRED PREDICT ; VAR    ASVABGI ASVABAR
ASVABGS ASVABMC ENTRYAGE ASVABAD;
DATA TABLE;SET TWOPRED;
PRDICTOR=(LOSM1HAT+LOSM2HAT)/2;
PROC SORT;        BY PRDICTOR;
TITLE  SORTING BY PRDICTOR FOR OTHER SH MEN;
DATA STORE;SET TABLE;FILE FT34F001;
PUT (FORYEAR PRDICTOR THRYEAR RACE) (10.7);
/*
```

TABLE XXIX

Program Listing of FORTRAN Program for Cut-offs

```
      REAL PASS(4000),SCORE(3,4000),PASST(4000),
     *     RETURN(3,4000),TOTRET(3,3),NPLOT(3),
     *     SUCALL(3),RACE(4000),SRALL(3),UTILSP(2),
     *     UTILSF(2),SCOMAX(3,2),TOTSC(3,3),
     *     S(4000),R(4000),TMS(3),TMR(3)
      DATA UTILSP/-9244.,18488.0/
      DATA UTILSF/14495.0,-14495./
      SINGLE=3.
      D=55.0
      E=25.0
      DO 100 J=1,3
      I=1
      PASSA=0.0
      PASSB=0.0
      FAILA=0.0
      FAILB=0.0
      PASS(I)=0.0
      SCORE(J,I)=30.0
      RETURN(J,I)=0.0
      RETURN(J,I+1)=0.0
      RACE(I)=0.0
   10 FORMAT (4F10.7)
   20 I=I+1
      READ(J,10,END=50) PASS(I),SCORE(J,I),
     *                  PASST(I),RACE(I)
      PASSA=PASSA+PASS(I)
      GO TO 20
   50 CONTINUE
      NPLOT(J)=I-1
      GTOTAL=I-2
      FAILA=GTOTAL-PASSA
      X=PASSA*UTILSP(2)+PASSB*UTILSP(1)
      X=X+FAILA*UTILSF(2)+FAILB*UTILSF(1)
      BASEUT=X/GTOTAL
      N=NPLOT(J)
      WRITE(6,109)J,GTOTAL,BASEUT,N
  109 FORMAT(I4,F10.3,F10.2,I8)
      DO 40 I=2,N
      IF (PASS(I)) 31,31,32
   31 FAILB=FAILB+1
      FAILA=FAILA-1
      GO TO 33
   32 PASSB=PASSB+1
      PASSA=PASSA-1
   33 X=PASSA*UTILSP(2)+PASSB*UTILSP(1)
      X=X+FAILA*UTILSF(2)+FAILB*UTILSF(1)
      X=X/GTOTAL
      IF (BASEUT.NE.0.0) GO TO 77
      BASEUT=X
      IF (X.NE.0.)GO TO 77
      RETURN(J,I+1)=0.
      GO TO 25
   77 RETURN(J,I+1)=100*(X-BASEUT)/(BASEUT)
      IF(BASEUT.GT.0.0)GO TO 25
      RETURN(J,I+1)=-RETURN(J,I+1)
   25 IF (I.EQ.2) GO TO 34
      IF (RETURN(J,I+1).LT.SCOMAX(J,2)) GO TO 38
   34 SCOMAX(J,2)=RETURN(J,I)
      SCOMAX(J,1)=SCORE(J,I)
      SRALL(J)=((PASSA+FAILA)/GTOTAL)*100
      IF (SRALL(J).GT.0.0)GO TO 53
      SUCALL(J)=0.0
   53 SUCALL(J)=((PASSA)/(PASSA+FAILA))*100
   38 IF (RETURN(J,I).GT.-20.0) GO TO 40
```

```
      RETURN(J,I)=-20.0
 40   CONTINUE
      TOTSC(J,1)=SCOMAX(J,1)
      TOTSC(J,2)=SCOMAX(J,1)
      TOTSC(J,3)=D
      TOTRET(J,1)=-20.0
      TOTRET(J,2)=SCOMAX(J,2)
      TOTRET(J,3)=E
      E=E-4.5
 100  CONTINUE
      CALL TEK618
      CALL BLOWUP  (1.5)
      CALL MX1ALF  ('STANDARD','/')
      CALL MX2ALF  ('L/CSTD','+')
      CALL YTICKS  (5)
      CALL XTICKS  (5)
      CALL PAGE    (7.5,6.0)
      CALL AREA2D  (6.5,4.3)
      CALL XNAME  ('C+UT-/O+FF ON /P+REDICTOR
     *                          /E+QUATION',36)
      CALL YNAME  ('P+ERCENTAGE /C+HANGE
     *                  FROM /B+ASE /R+ATE',39)
      CALL HEADIN  ('SH M+EN ON /LOS
     *                       C+RITERION',26,-1.,3)
      CALL HEADIN  ('R+ACE /S+PECIFIC
     *                   /S+ELECTION',28,-1.,3)
      CALL HEADIN  ('F+OR /A+LL /SH M+EN',19,-1.,3)
      CALL GRAF  (30,5,70,-20,10,30)
      DO 200 J=1,3
      N=NPLOT(J)
      DO 150 I=1,N
      S(I)=SCORE(J,I)
      R(I)=RETURN(J,I)
      IF (I.GT.3) GO TO 150
      TMS(I)=TOTSC(J,I)
      TMR(I)=TOTRET(J,I)
 150  CONTINUE
      IF (J.EQ.1) GO TO 160
      IF (J.EQ.2) GO TO 155
      CALL DASH
      GO TO 160
 155  CALL DOT
 160  CALL CURVE  (S,R,N,0)
      CALL CURVE  (TMS,TMR,3,0)
      CALL HEIGHT  (.09)
      CALL RLMESS  ('M+AXIMUM /R+ETURN: ',19,
     *                      TOTSC(J,3),TOTRET(J,3))
      IF (J.EQ.1)  GO TO 110
      IF (J.EQ.2)  GO TO 120
      CALL RLMESS  ('O+THERS',7,'ABUT','ABUT')
      GO TO 130
 120  CALL RLMESS  ('B+LACKS',7,'ABUT','ABUT')
      GO TO 130
 110  CALL RLMESS  ('W+HITES',7,'ABUT','ABUT')
 130  CALL RLREAL  (SCOMAX(J,1),2,
                 TOTSC(J,3)+2.,TOTRET(J,3)-2.)
      CALL RLMESS  (',',1,'ABUT','ABUT')
      CALL RLREAL  (SCOMAX(J,2),2,'ABUT','ABUT')
      CALL RLMESS  ('%',1,'ABUT','ABUT')
 200  CONTINUE
      D=5.0
      E=-5.0
      CALL RLMESS  ('SELECTION RATIOS:',17,55.0,D)
      CALL RLMESS  ('PERCENTAGE SUCCESSFUL:',
                                   22,55.0,E)
      DO 300 J=1,3
      D=D-2.0
      E=E-2.0
      IF (J.EQ.1)  GO TO 310
```

```
      IF (J.EQ.2) GO TO 320
      CALL RLMESS ('   O+THER /G+ROUP :',18,55.0,D)
      CALL REALNO (SRALL(J),1,'ABUT','ABUT')
      CALL RLMESS ('%',1,'ABUT','ABUT')
      CALL RLMESS ('   O+THER /G+ROUP :',18,55.0,E)
      CALL REALNO (SUCALL(J),1,'ABUT','ABUT')
      CALL RLMESS ('%',1,'ABUT','ABUT')
      GO TO 300
  320 CALL RLMESS ('   B+LACK /G+ROUP :',18,55.0,D)
      CALL REALNO (SRALL(J),1,'ABUT','ABUT')
      CALL RLMESS ('%',1,'ABUT','ABUT')
      CALL RLMESS ('   B+LACK /G+ROUP :',18,55.0,E)
      CALL REALNO (SUCALL(J),1,'ABUT','ABUT')
      CALL RLMESS ('%',1,'ABUT','ABUT')
      GO TO 300
  310 CALL RLMESS ('   W+HITE /G+ROUP :',18,55.0,D)
      CALL REALNO (SRALL(J),1,'ABUT','ABUT')
      CALL RLMESS ('%',1,'ABUT','ABUT')
      CALL RLMESS ('   W+HITE /G+ROUP :',18,55.0,E)
      CALL REALNO (SUCALL(J),1,'ABUT','ABUT')
      CALL RLMESS ('%',1,'ABUT','ABUT')
  300 CONTINUE
      CALL ENDPL (C)
      CALL DONEPL
      STOP
      END
```

APPENDIX B

STEPWISE REGRESSION RESULTS FOR RATINGS

This appendix gives the results of the stepwise regressions. They are presented in the form of rating specific tables. As these tables were constructed directly from computer output the variables all have the SAS names that were used in the analysis. Therefore, interested readers are directed to the relevant sections of the preceding SAS program listings to determine exactly which variables are listed. In all cases the variables and their coefficient signs correspond to the tables in the body of the thesis which lists the stepwise regression results.

TABLE XXX

All SH Stepwise Regression Results

SH Whites on Length of Service

	DF	SUM OF SQU	F	PROB>F
REGRESSION	8	16236.804	11.08	0.0001
ERROR	1297	237637.401		
TOTAL	1305	253874.205		

	B VALUE	STD E	F	PROB>F
INTERCEPT	26.6701644			
AFQTGRPS	-0.9209895	0.3761	5.99	0.0145
ASVAEMC	-0.2688282	0.1360	3.90	0.0484
ASVAEAI	-0.1703922	0.0978	3.03	0.0818
ASVAECA	-0.1555049	0.0917	2.87	0.0903
ENTRPAYG	1.7093734	0.8128	4.42	0.0357
MRTSTAT1	5.6677671	0.8366	45.89	0.0001
SCREEN	0.2485786	0.0607	16.73	0.0001
DEPNDNTS	-3.8832762	2.1708	3.20	0.0739

SH Whites on GOODGUY

	DF	SUM OF SQU	F	PROB>F
REGRESSION	4	3022.5064	12.22	0.0001
ERROR	1301	80462.1030		
TOTAL	1305	83484.6094		

	B VALUE	STD E	F	PROB>F
INTERCEPT	1.10711196			
ASVAEMK	0.08789585	0.0561	2.45	0.1174
ASVAECE	-0.14122207	0.0493	8.18	0.0043
CHYEC	1.36078242	0.2466	30.43	0.0001
MRTSTAT1	1.09421802	0.4740	5.33	0.0211

SH Whites on BADGUY

	DF	SUM OF SQU	F	PROB>F
REGRESSION	7	4135.7682	12.74	0.0001
ERROR	1298	60204.2011		
TOTAL	1305	64339.9693		

	B VALUE	STD E	F	PROB>F
INTERCEPT	-1.74204431			
ASVAENO	0.03938113	0.0225	3.04	0.0816
ASVAEWK	-0.12244891	0.0420	8.47	0.0037
ASVAECE	-0.09038137	0.0428	4.45	0.0352
CHYEC	1.20950502	0.2603	21.58	0.0001
ENTRPAYG	0.90472456	0.4131	4.80	0.0287
MRTSTAT1	1.13491198	0.4115	7.61	0.0059
SCREEN	0.08475720	0.0359	5.56	0.0186

SH Blacks on Length of Service

	DF	SUM OF SQU	F	PROB>F
REGRESSION	5	9203.0966	11.26	0.0001
ERROR	552	90267.1489		
TOTAL	557	99470.2455		

	B VALUE	STD E	F	PROB>F
INTERCEPT	32.7376235			
AFQTGRPS	-1.9561139	0.4563	18.37	0.0001
ASVABCA	-0.4280135	0.1625	6.93	0.0087
ASVAECC	0.3140531	0.1357	5.35	0.0211
MRTSTAT1	6.2469389	1.1323	30.43	0.0001
SCREEN	0.2155208	0.0872	6.10	0.0138

SH Blacks on GOODGUY

	DF	SUM OF SQU	F	PROB>F
REGRESSION	4	621.0247	3.34	0.0102
ERROR	553	25714.4591		
TOTAL	557	26335.4838		

	B VALUE	STD E	F	PROB>F
INTERCEPT	11.4082688			
AFQTGRPS	-0.6612232	0.2428	7.41	0.0067

ASVABCC	0.0992502	0.0590	2.83	0.0931
MRTSTAT1	1.1435432	0.6035	3.59	0.0586
SCREEN	0.0855462	0.0460	3.45	0.0639

SH Blacks on BADGUY

	DF	SUM OF SQU	F	PROB>F
REGRESSION	4	900.1514	6.76	0.0001
ERROR	553	18399.6692		
TOTAL	557	19299.8207		
	B VALUE	STD E	F	PROB>F
INTERCEPT	11.1998625			
ASVABAR	-0.2490178	0.0778	10.23	0.0015
ASVABCC	0.1040409	0.0497	4.38	0.0369
ENTRYAGE	-0.1948975	0.1156	2.84	0.0926
CHYEC	1.1681247	0.3167	13.60	0.0002

SH Others on Length of Service

	DF	SUM OF SQU	F	PROB>F
REGRESSION	6	8636.5314	11.36	0.0001
ERROR	149	18874.4685		
TOTAL	155	27511.0000		
	B VALUE	STD E	F	PROB>F
INTERCEPT	54.9934418			
ASVABGI	-1.1711530	0.3280	12.74	0.0005
ASVABAD	-0.3308399	0.1989	2.77	0.0984
ASVABAR	-0.8798285	0.2952	8.88	0.0034
ASVABMC	-0.7107025	0.3736	3.62	0.0591
ASVABGS	1.1684002	0.3674	10.11	0.0018
ENTRYAGE	0.6142226	0.3009	4.17	0.0430

SH Others on GOODGUY

	DF	SUM OF SQU	F	PROB>F
REGRESSION	5	1050.2840	5.95	0.0001
ERROR	150	5293.3056		
TOTAL	155	6343.5897		
	B VALUE	STD E	F	PROB>F
INTERCEPT	8.68073217			
AFQTPCNT	-0.09486851	0.0286	10.93	0.0012
ASVABEI	-0.16364685	0.1001	2.68	0.1039
ASVABAI	-0.28815095	0.1548	3.46	0.0647
ASVABCC	0.38797354	0.1097	12.49	0.0005
SCREEN	0.14637842	0.0801	3.33	0.0698

SH Others on BADGUY

	DF	SUM OF SQU	F	PROB>F
REGRESSION	5	644.5853	4.89	0.0004
ERROR	150	3952.8505		
TOTAL	155	4597.4358		
	B VALUE	STD E	F	PROB>F
INTERCEPT	6.77090112			
AFQTPCNT	-0.16194490	0.0584	7.67	0.0063
AFQTGRPS	1.31562114	0.7261	3.28	0.0720
ASVABAI	-0.28525384	0.1288	4.90	0.0284
ASVABCC	0.23515378	0.0952	6.09	0.0147
CHYEC	0.98156281	0.4234	5.37	0.0218

TABLE XXXI

All PN Stepwise Regression Results

PN White Men on Length of Service

	DF	SUM OF SQU	F	PROB>F
REGRESSION	9	23114.535	14.78	0.0001
ERROR	1216	211247.111		
TOTAL	1225	234361.647		

	B VALUE	STD E	F	PROB>F
INTERCEPT	30.6886800			
ASVABWK	-0.3907675	0.0918	18.10	0.0001
ASVABAR	-0.2914101	0.1277	5.20	0.0227
ASVABEI	0.2812350	0.1048	7.19	0.0074
ASVABSI	-0.2984347	0.1175	6.44	0.0113
ASVABCE	0.1368159	0.0849	2.59	0.1076
ENTRYAGE	0.3694307	0.1831	4.07	0.0439
CHYEC	-1.0215411	0.4358	5.49	0.0193
MRTSTAT1	7.6648584	0.7969	92.49	0.0001
SCREEN	0.2812599	0.0781	12.96	0.0003

PN White Men on GOODGUY

	DF	SUM OF SQU	F	PROB>F
REGRESSION	8	2013.1708	5.33	0.0001
ERROR	1217	57483.8112		
TOTAL	1225	59496.9820		

	B VALUE	STD E	F	PROB>F
INTERCEPT	10.4760049			
ASVABWK	-0.1122275	0.0443	6.41	0.0115
ASVABAR	-0.2187153	0.0682	10.28	0.0014
ASVABMC	0.0974479	0.0637	2.34	0.1267
ASVABCE	0.1244157	0.0424	8.58	0.0035
CHYEC	-0.3040941	0.1929	2.48	0.1154
MRTSTAT1	0.9357586	0.4190	4.99	0.0257
SCREEN	0.1642128	0.0406	16.33	0.0001
DEPNDNTS	1.5133497	0.8221	3.39	0.0659

PN White Men on BADGUY

	DF	SUM OF SQU	F	PROB>F
REGRESSION	4	626.2016	5.70	0.0002
ERROR	1221	33550.0625		
TOTAL	1225	34176.2642		

	B VALUE	STD E	F	PROB>F
INTERCEPT	12.9046054			
AFCTGRPS	-0.5259052	0.1598	10.83	0.0010
ASVABEI	0.0801835	0.0373	4.60	0.0321
SCREEN	0.1092251	0.0268	16.54	0.0001
DEPNDNTS	1.1273538	0.6069	3.45	0.0635

PN Black Men on Length of Service

	DF	SUM OF SQU	F	PROB>F
REGRESSION	6	6750.8674	8.31	0.0001
ERROR	276	37349.7756		
TOTAL	282	44100.6431		

	B VALUE	STD E	F	PROB>F
INTERCEPT	48.8556818			
AFCTPCNT	-0.1731397	0.0417	17.17	0.0001
ASVABNO	-0.1314724	0.0864	2.31	0.1295
ASVABCC	-0.3337261	0.1623	4.23	0.0408
CHYEC	-2.4543496	0.5735	18.31	0.0001
MRTSTAT1	5.8937313	1.4492	16.54	0.0001
SCREEN	0.4812862	0.1509	10.17	0.0016

PN Black Men on GOODGUY

	DF	SUM OF SQU	F	PROB>F
REGRESSION	7	950.4489	3.68	0.0009
ERROR	275	10156.2648		
TOTAL	282	11106.7137		

	B VALUE	STD E	F	PROB>F
INTERCEPT	24.9538629			
ASVABGI	0.3102658	0.1504	4.26	0.0401
ASVABNO	-0.1398792	0.0472	8.77	0.0033
ASVABMK	0.1769179	0.1037	2.91	0.0894
ASVABEI	0.1726309	0.0956	3.26	0.0721
ASVABSI	-0.3085155	0.1207	6.53	0.0112
ASVABCE	-0.1521711	0.0810	3.52	0.0615
CHYEC	-0.5306986	0.2566	4.27	0.0396

PN Black Men on BADGUY

	DF	SUM OF SQU	F	PROB>F
REGRESSION	4	390.1637	5.31	0.0004
ERROR	278	5106.6559		
TOTAL	282	5496.8197		

	B VALUE	STD E	F	PROB>F
INTERCEPT	27.3597073			
ASVABSI	-0.2845607	0.0866	10.80	0.0011
ASVABAI	0.2497785	0.0890	7.87	0.0054
ENTRYAGE	-0.3561344	0.1018	12.22	0.0006
ENTRPAYG	0.5638719	0.3534	2.54	0.1118

PN Other Men on Length of Service

	DF	SUM OF SQU	F	PROB>F
REGRESSION	3	3425.1745	9.05	0.0001
ERROR	87	10976.1221		
TOTAL	90	14401.2967		

	B VALUE	STD E	F	PROB>F
INTERCEPT	37.3172602			
ASVABGI	-0.9104034	0.3173	8.23	0.0052
CHYEC	1.8625773	0.8067	5.33	0.0233
DEPNDNTS	-12.7447520	3.9740	10.28	0.0019

PN Other Men on GOODGUY

	DF	SUM OF SQU	F	PROB>F
REGRESSION	5	774.4225	4.99	0.0005
ERROR	85	2636.5664		
TOTAL	90	3410.9890		

	B VALUE	STD E	F	PROB>F
INTERCEPT	0.52074631			
AFQTPCNT	-0.12040574	0.0395	9.28	0.0031
ASVABAR	0.30390695	0.2059	2.18	0.1437
ASVABCC	0.27792854	0.1166	5.68	0.0194
CHYEC	1.33872861	0.4117	10.57	0.0016
DEPNDNTS	-4.69898059	1.9942	5.55	0.0208

PN Other Men on BADGUY

	DF	SUM OF SQU	F	PROB>F
REGRESSION	4	423.9178	5.34	0.0007
ERROR	86	1705.7524		
TOTAL	90	2129.6703		

	B VALUE	STD E	F	PROB>F
INTERCEPT	5.84043938			
AFQTPCNT	-0.04974738	0.0219	5.16	0.0257
ASVABSI	0.37905240	0.1304	8.44	0.0047
CHYEC	1.08168182	0.3278	10.89	0.0014
DEPNDNTS	-3.35506552	1.5848	4.48	0.0372

PN White Women on Length of Service

	DF	SUM OF SQU	F	PROB>F
REGRESSION	6	7047.9736	6.31	0.0001
ERROR	472	87924.6777		
TOTAL	478	94972.6513		

	B VALUE	STD E	F	PROB>F
INTERCEPT	95.2187469			
ASVABAD	-0.3924025	0.1608	5.95	0.0151
ASVABWK	-0.3429857	0.1734	3.91	0.0485
ASVABAR	-0.4423221	0.2133	4.30	0.0387
ASVABCE	0.5541282	0.1689	10.75	0.0011

CHYEC	-2.7667462	0.8557	10.45	0.0013
ENTRPAYG	1.7744842	1.1070	2.57	0.1096

PN White Women on GOODGUY

	DF	SUM OF SQU	F	PROB>F
REGRESSION	3	971.0785	5.62	0.0010
ERROR	475	27338.7335		
TOTAL	478	28309.8121		

	B VALUE	STD E	F	PROB>F
INTERCEPT	29.1338691			
ASVABAD	-0.2337727	0.0893	6.85	0.0092
ASVABWK	-0.1576776	0.0935	2.84	0.0924
ASVABAR	-0.2607376	0.1169	4.97	0.0262

PN White Women on BADGUY

	DF	SUM OF SQU	F	PROB>F
REGRESSION	5	576.7302	5.14	0.0002
ERROR	473	10619.5119		
TOTAL	478	11196.2421		

	B VALUE	STD E	F	PROB>F
INTERCEPT	19.4431146			
ASVABAD	-0.1212187	0.0556	4.75	0.0299
ASVABAR	-0.2148259	0.0733	8.57	0.0036
ASVABSP	0.1078479	0.0632	2.91	0.0887
CHYEC	0.4943182	0.2339	4.47	0.0351
DEPNDNTS	-2.6235293	0.9765	7.22	0.0075

TABLE XXXII

All AT Stepwise Regression Results

AT White Men on Length of Service

	DF	SUM OF SQU	F	PROB>F
REGRESSION	12	37330.665	29.55	0.0001
ERROR	3327	350277.988		
TOTAL	3339	387608.653		
	B VALUE	STD E	F	PROB>F
INTERCEPT	49.7303064			
AFQTPCNT	0.0847810	0.0278	9.24	0.0024
AFQTGRPS	-0.8119479	0.5030	2.60	0.1066
ASVABNO	-0.0788686	0.0243	10.46	0.0012
ASVABSP	-0.1110189	0.0670	2.74	0.0977
ASVABMK	-0.1183909	0.0684	2.99	0.0839
ASVABGS	-0.1355446	0.0712	3.61	0.0574
ASVABCA	-0.2048483	0.0586	12.19	0.0005
ASVABCE	0.1205116	0.0453	7.07	0.0079
CHYEC	-0.7428230	0.2858	6.75	0.0094
ENTRPAYG	2.1286892	0.1946	119.62	0.0001
MRTSTAT1	4.6121976	0.3735	152.47	0.0001
SCREEN	0.0957458	0.0406	5.56	0.0185

AT White Men on GOODGUY

	DF	SUM OF SQU	F	PROB>F
REGRESSION	3	456.839	3.27	0.0202
ERROR	3336	155283.609		
TOTAL	3339	155740.449		
	B VALUE	STD E	F	PROB>F
INTERCEPT	13.8263146			
CHYEC	0.3864369	0.1663	5.40	0.0202
ENTRPAYG	-0.1868022	0.1270	2.16	0.1415
MRTSTAT1	0.5036740	0.2450	4.23	0.0399

AT White Men on BADGUY

	DF	SUM OF SQU	F	PROB>F
REGRESSION	6	4061.6007	32.17	0.0001
ERROR	3333	70130.5848		
TOTAL	3339	74192.1856		
	B VALUE	STD E	F	PROB>F
INTERCEPT	12.1536983			
AFQTGRPS	-0.3680884	0.1027	12.82	0.0003
ENTRYAGE	0.0849608	0.0440	3.71	0.0541
CHYEC	0.4063458	0.1381	8.65	0.0033
ENTRPAYG	-0.9321898	0.0862	116.75	0.0001
MRTSTAT1	-0.4857506	0.1669	8.47	0.0036
SCREEN	0.0891893	0.0183	23.64	0.0001

AT Black Men on Length of Service

	DF	SUM OF SQU	F	PROB>F
REGRESSION	5	3777.9917	7.66	0.0001
ERROR	167	16480.0660		
TOTAL	172	20258.0578		
	B VALUE	STD E	F	PROB>F
INTERCEPT	76.0373893			
ASVABGS	-0.6729970	0.2328	8.35	0.0044
ASVABSI	-0.4349189	0.2312	3.54	0.0617
ASVABCA	0.3428155	0.1777	3.72	0.0555
CHYEC	-1.7461861	0.7264	5.78	0.0173
MRTSTAT1	4.5120091	1.5433	8.55	0.0039

AT Black Men on GOODGUY

	DF	SUM OF SQU	F	PROB>F
REGRESSION	5	1239.3404	7.99	0.0001
ERROR	167	5181.4687		
TOTAL	172	6420.8092		
	B VALUE	STD E	F	PROB>F

INTERCEPT	47.2311662			
AFQTPCNT	0.2447246	0.0711	11.82	0.0007
AFQTGRPS	-4.9159979	1.2330	15.89	0.0001
ASVABGS	-0.2445383	0.1597	2.34	0.1277
ASVABCA	0.1803334	0.1012	3.17	0.0768
CHYEC	-1.0599272	0.4173	6.45	0.0120

AT Black Men on BADGUY

	DF	SUM OF SQU	F	PROB>F
REGRESSION	6	461.4027	5.82	0.0001
ERROR	166	2191.7764		
TOTAL	172	2653.1791		
	B VALUE	STD E	F	PROB>F
INTERCEPT	29.5616565			
AFQTPCNT	0.1124665	0.0452	6.17	0.0140
AFQTGRPS	-3.0880771	0.7941	15.12	0.0001
ASVABAR	0.3320927	0.1154	8.28	0.0045
ASVABMK	-0.1282669	0.0869	2.18	0.1421
ASVABAI	0.1147313	0.0618	3.44	0.0653
ENTRPAYG	-0.6578062	0.3131	4.41	0.0372

AT Other Men on Length of Service

	DF	SUM OF SQU	F	PROB>F
REGRESSION	11	5054.3481	5.96	0.0001
ERROR	98	7557.6518		
TOTAL	109	12612.0000		
	B VALUE	STD E	F	PROB>F
INTERCEPT	-4.67799655			
ASVABEI	0.60422952	0.2241	7.27	0.0083
ASVABAI	-0.99876644	0.2561	15.21	0.0002
ASVABCM	0.77440682	0.1924	16.20	0.0001
ENTRYAGE	1.40536754	0.4478	9.85	0.0022
SCREEN	0.71935965	0.1641	19.22	0.0001

AT Other Men on GOODGUY

	DF	SUM OF SQU	F	PROB>F
REGRESSION	10	2295.9445	6.94	0.0001
ERROR	99	3275.8736		
TOTAL	109	5571.3181		
	B VALUE	STD E	F	PROB>F
INTERCEPT	8.85750116			
AFQTPCNT	-0.18425571	0.0504	13.35	0.0004
ASVABAD	0.51458563	0.1369	14.11	0.0003
ASVABWK	0.86805391	0.1648	27.72	0.0001
ASVABEI	-0.33035220	0.1353	5.96	0.0164
SCREEN	0.34213980	0.1089	9.86	0.0022

AT Other Men on BADGUY

	DF	SUM OF SQU	F	PROB>F
REGRESSION	12	848.2633	6.55	0.0001
ERROR	97	1047.1911		
TOTAL	109	1895.4545		
	B VALUE	STD E	F	PROB>F
INTERCEPT	13.4256325			
ASVABMK	0.3560198	0.1186	9.00	0.0034
ASVABEI	-0.3764082	0.0825	20.77	0.0001
ASVABAI	-0.2365056	0.1077	4.81	0.0306
SCREEN	0.1428342	0.0606	5.55	0.0205

AT White Women on Length of Service

	DF	SUM OF SQU	F	PROB>F
REGRESSION	4	3820.0507	5.05	0.0006
ERROR	235	44438.4450		
TOTAL	239	48258.4958		
	B VALUE	STD E	F	PROB>F
INTERCEPT	39.2507434			
ASVABEI	0.7791086	0.2424	10.33	0.0015
ASVABCM	0.3521168	0.2423	2.11	0.1476
ASVABCE	-0.5763499	0.2557	5.08	0.0252

MRTSTAT1	-5.1232380	1.7929	8.17	0.0047

AT White Women on GOODGUY

	DF	SUM OF SQU	F	PROB>F
REGRESSION	3	545.1452	3.93	0.0092
ERROR	239	11038.3938		
TOTAL	242	11583.5390		

	B VALUE	STD E	F	PROB>F
INTERCEPT	12.9788322			
ASVABEI	0.2686742	0.1179	5.19	0.0236
ASVABCM	0.2374158	0.1197	3.93	0.0485
ASVABCE	-0.3717212	0.1260	8.70	0.0035

AT White Women on BADGUY

	DF	SUM OF SQU	F	PROB>F
REGRESSION	6	486.3025	5.22	0.0001
ERROR	236	3667.6069		
TOTAL	242	4153.9094		

	B VALUE	STD E	F	PROB>F
INTERCEPT	16.5036012			
ASVABCE	-0.1729768	0.0651	7.04	0.0085
ASVABCC	0.1377274	0.0555	6.14	0.0139
ENTRYAGE	0.1883614	0.0896	4.42	0.0366
ENTRRAYG	-0.8121685	0.2866	8.03	0.0050
MRTSTAT1	1.3441319	0.5330	6.36	0.0123
DEPNDNTS	-3.3888684	1.0103	11.25	0.0009

LIST OF REFERENCES

1. Assessment Group. Billet Cost Model. Santa Monica. 1982.

2. Barrett, G. V., Phillips, J. S. & Alexander, R. A. Concurrent and Predictive Validity Designs: A Critical Re-analysis. Journal of Applied Psychology, 1981, 66, 1-6.

3. Balis, E. & Clay-Mendez, P. Replacement Costs for Navy First Term Personnel by Rating. CNA Report 82-1357. Center for Naval Analyses: Alexandria. September 1982.

4. Butler R. Imputation of a Sailor's Marginal Product: An Application of the Theory of Human Capital. RD 126. Assessment Group: Santa Monica. July 1982.

5. Campbell, J. P. Psychometric Theory. In Dunnette, M. D. Handbook of Industrial and Organizational Psychology. Chicago: Rand McNally, 1976.

6. Carroll, J. B., & Horn, J. L. On the Scientific Basis of Ability Testing. American Psychologist, 1981, 36, 1012-1020.

7. Drenth, P. Theory and Methods of Selection. In Warr, P. Psychology at Work. London: Penguin, 1971.

8. Dunnette, M. D. Handbook of Industrial and Organizational Psychology. Chicago: Rand McNally, 1976.

9. Ghiselli, E. E. The Validity of Occupational Tests. New York: Wiley & Sons, 1966.

10. Guion, R. M. Recruiting, Selection and Job Placement. In Dunnette, M. D. Handbook of Industrial and Organizational Psychology. Chicago: Rand McNally, 1976.

11. Guion, R. M. & Campbell, J. P. A Note on Concurrent and Predictive Validity Designs: A Critical Analysis. Journal of Applied Psychology, 1982, 67, 239-244.

12. Hammer, T. H. & Landau, J. Methodological Issues in the Use of Absence Data. Journal of Applied Psychology, 1981, 66, 574-581.

13. Hsu, L. M. Estimation of Relative Validities of Selection Tests from Information Commonly Available in Presence of Direct and Indirect Range Restriction. Journal of Applied Psychology, 1982, 67, 509-511.

14. Jensen, A. R. Genetics and Education. London: Methuen, 1973.

15. Jensen, A. R. Bias in Mental Testing. New York: Free Press, 1980.

16. Johnson, I.D. Abstracts of Reported Military Studies on Manpower Attraction, Retention and Attrition. Technical Note 2/82. 1 Psychological Research Unit: Canberra. February 1982.

17. Landy, F. J. & Farr, J. L. Performance Rating Psychological Bulletin, 1980, 87, 72-107.

18. Landy, F. J., Farr, J. L., & Jacobs, R. R. Utility Concepts in Performance Measurement. Organizational Behaviour and Human Performance, 1982, 30, 15-40.

19. Lurie, P.E. Relating Enlistment Standards to Job Performance: a Pilot Study for Two Navy Ratings. CNA Report 81-0048. Center for Naval Analyses: Alexandria. January 1981.

20. McNemar, Q. Psychological Statistics. New York: John Wiley & Sons. 1963.

21. O'Gorman, J.G. Predicting Effectiveness for General Enlistment to the Australian Army. Report No 9/72. 1 Psychological Research Unit: Melbourne. October 1972.

22. Reilly, R. R. & Chao, G. T. Validity and Fairness of Some Alternative Employee Selection Procedures. Personnel Psychology, 1982, 35, 1-62.

23. Rowland, K. M., London, M., Ferris, G. R. & Sherman, J. L. Current Issues in Personnel Management. New York: Allyn & Bacon, 1980.

24. Sands, W.A. Screening Male Applicants for Navy Enlistment. TR 77-34. Navy Personnel Research and Development Centre: San Diego. June 1977.

25. Schmidt, F. L. & Hunter, J. E. Employment Testing: Old Theories and New Research. American Psychologist, 1981, 36, 1128-1137.

26. Smith, P. C. Behaviours, Results and Organizational Effectiveness: The Problem. In Dunnette, M. D. _Handbook of Industrial and Organizational Psychology_. Chicago: Rand McNally, 1976.

27. SAS Institute. _Statistical Analysis User's Guide_. North Carolina: SAS Institute, 1979.

28. Tiffin, J. & McCormack, E. J. _Industrial Psychology_. London: Prentice-Hall, 1970.

29. _US Navy Publication_. Enlisted Career Guide 1981-82. Navy Recruiting Command: Washington. 1981.

30. Weiss, D. J. Multivariate Procedures. In Dunnette, M. D. _Handbook of Industrial and Organizational Psychology_. Chicago: Rand McNally, 1976.

INITIAL DISTRIBUTION LIST

		No. Copies
1.	Dudley Knox Library Code 0142 Naval Postgraduate School Monterey, CA 93940	2
2.	Defense Technical Information Center Cameron Station Alexandria, VA 22314	2
3.	Defence Library Campbell Park Offices Canberra, ACT 2600 Australia	2
4.	Department Chairman, Code 54 Department of Administrative Sciences Naval Postgraduate School Monterey, CA 93940	1
5.	Professor Richard S. Elster, Code 54Ea Department of Administrative Sciences Naval Postgraduate School Monterey, CA 93940	3
6.	Professor John A. Senger, Code 54Se Department of Administrative Sciences Naval Postgraduate School Monterey, CA 93940	3
7.	Professor W. McGavey, Code 54 Department of Administrative Sciences Naval Postgraduate School Monterey, CA 93940	2
8.	Professor Paul Hoffman, Code 54 Department of Administrative Sciences Naval Postgraduate School Monterey, CA 93940	1
9.	Professor George Thomas, Code 54Te Department of Administrative Sciences Naval Postgraduate School Monterey, CA 93940	1
10.	Major Kelvin W. Nesbitt, AA Directorate of Personnel Plans-Army Russell Offices Canberra, ACT 2600 Australia	2

11. Dr. Martin Wiskoff, Code 12 1
Navy Personnel R&D Center
San Diego, CA 92152

12. Dr. Robert Lockman 1
Center for Naval Analyses
Bureauregard St.
Alexandria, VA 22314

13. Deputy Chief of Naval Operations (M,P,T) 1
Attn: OP-11, -12, -13
Arlington Annex
Washington, D.C. 20370

14. Major Frank Pinch 1
CFPARU
Suite 600, 4900 Yonge St.
Willowdale, Ont M2N-6B7

15. MPTA Library, Code 36 1
Naval Postgraduate School
Monterey, CA 93940

16. Commanding Officer 1
1 Psychological Research Unit
Campbell Park Offices
Canberra ACT 2600 Australia

END

DATE
FILMED

9 — 83

DTIC